

Tesis Doctoral

LAS FACETAS DE LOS BIG FIVE Y LA PREDICCIÓN DEL DESEMPEÑO

Doctoral Dissertation

THE BIG FIVE FACETS AND PERFORMANCE PREDICTION

Beata Chorągwicka

Director: Prof. Jesús F. Salgado

Santiago de Compostela, Junio 2010

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

JESÚS F. SALGADO VELO, Catedrático de Psicología Social de la Universidad de Santiago de Compostela y director de la tesis doctoral titulada "Las facetas de los Big Five y la predicción del desempeño", presentada por Dª Beata Choragwicka, y como trámite perceptivo para su aceptación y posterior defensa pública, INFORMA:

Que la mencionada tesis es un trabajo de investigación original y que fue proyectada, desarrollada y redactada bajo mi continua supervisión. A mi juicio, dicha tesis reúne todas las características técnicas, metodológicas y científicas para ser leída y defendida públicamente y merece la máxima valoración en cuanto a la seriedad de planteamiento, la actualidad del enfoque, el rigor metodológico y la calidad e interés de las aportaciones efectuadas al conocimiento del objeto de estudio.

Santiago de Compostela a 14 de junio 2010

Fdo: Jesús F. Salgado

ACKNOWLEDGMENTS

This doctoral dissertation was written while I was working as a Research and Teaching Assistant at the Department of Social and General Psychology and Methodology at the University of Santiago de Compostela in Spain with a grant from the Spanish Ministry of Education (FPU, AP2005-3807). I would like to thank both the Spanish Ministry of Education and the University of Santiago de Compostela for their assistance and for making my dream of pursuing PhD studies in the HR field possible.

Apart from this institutional support, I am well aware that a large number of people were vital in making this dissertation possible. This work is actually the result of their help, their encouragement, their love, their support, their time and their patience. No words would be enough to express my gratefulness to all of them.

First and foremost I would like to thank to my advisor, Professor Jesús Salgado, who has certainly made a huge impact on me during these five years, and whom I admire for his professionalism and sparkling intellect, and for his dedication to perfection. He introduced me to the field and taught me all the nuts and bolts of doing research in W/O psychology, he believed in me when nobody else did and he supported me in every moment. He passed his enthusiasm for bandwidth-fidelity dilemma issues on to me and he made me believe I was capable of handling this ambitious project, and he was always available to share his expertise with me. He taught me to always strive for excellence in my work and showed me that even in science, white is not always white and black not always black. Thank you very much for all our fruitful discussions, for your guidance, for reading my chapters over and over again with enormous patience.

A second person who played a key role in my PhD is Professor Silvia Moscoso. She was the first to introduce me to the world of selection tools and procedures and to the issues of gender bias which later became my main research interests. She also supported me in my teaching tasks and always had time to respond to the hundreds of questions and queries I had. I also owe her a special debt of gratitude for her support

and inspiring insights about my Master of Philosophy (DEA) project and all the papers originating from this work. Your excellent advice and suggestions for improvements have helped me to improve the quality of my work and are the reason I have been able to come so far. Thank you for everything.

Special gratitude also goes to the management and staff of the Labor Relations School, where I was assigned to my teaching practice and where I have had my office during last four years. I could always count on your support for all of my projects and you really made me feel welcome. You made the place where I used to burn the midnight oil feel like my second home during all these years, thank you very much.

Furthermore, I would like to thank the following people: Filip Lievens and Frederik Anseel from Ghent University for your invaluable support during my research visit at Ghent University, for involving me in your projects, for sharing with me your expertise in publishing in scientific journals and for being my friends in a foreign country; Filip De Fruyt from Ghent University for sharing with me your opinions on current research on the Big Five Model and providing me with so many articles of interest; Neil Christiansen from Central Michigan University for providing me with some interesting and unpublished papers on the bandwidth-fidelity dilemma and for sharing with me your point of view on my research topic.

I would also like to thank my current and former colleagues: Yuritzi Bedolla and Natalia Ferreiro for helping me with the input of data, thanks to your companionship this tiring and monotonous job became fun; Claudia Silva for proofreading my chapters with a layperson´s eye; Ana Piñeiro for sharing the laughter and stress and for being there when I needed help; Pamela Alonso for reading the whole thesis more than just once, for our long discussions on bandwidth-fidelity topics and other related subjects, for your technical and moral support even late at night and on weekends and for showing me what it means to be a real colleague and friend, I am really glad I could share the last year of my PhD with you; Eveline Schollaert for infecting me with your enthusiasm and for sharing with me all the difficult and happy moments a PhD student can have; Bart Van de Ven for being a great office-mate in Ghent, for helping me

together with Eveline with my article database; to my other colleagues from Ghent and from Santiago: Mariolejn, Nele, Let, Tina, Toon, Bart, Greet, Kuhn, Hans, Gabriel, Gloria, Santos and Manuel for all the help you gave me when I was fighting with my “PhD-windmills”.

Last but not least, I should not forget the moral support of my family. It does not matter how many predictors I tested, you always produced the highest correlation coefficient with my success, and your reliability was always equal to 1.0. I want to thank my parents and my brother for the understanding they showed me when I decided to start a PhD program in a foreign country. I know how much you suffered all these years only seeing me twice a year or less, however, you never once complained and you always supported my decision. I love you so much. Finally, I want to thank my beloved husband Robert, who changed his whole life so that I could pursue my “PhD dream”. With you by my side there is no mountain I cannot climb. Thank you for your love, for your patience, for laughing and crying with me, for being always my friend and my shield. I could not have done it without you.

Santiago de Compostela, May 25th, 2010

“La tarea más importante en la vida de un individuo es dar la luz a sí mismo, convertirse en lo que es potencialmente. El producto más importante de su esfuerzo es su propia personalidad.”

Erich Fromm (1900-1980)
El hombre para si mismo (1947)

“Man's main task in life is to give birth to himself, to become what he potentially is. The most important product of his effort is his own personality”

Erich Fromm (1900-1980)
Man for Himself (1947)

ÍNDICE DE CONTENIDO

CAPÍTULO 3

TABLE OF CONTENTS

CHAPTER 3

CAPÍTULO 1

LA PERSONALIDAD Y LA PREDICCIÓN DEL DESEMPEÑO OCUPACIONAL

De acuerdo a Barrick, Mount y Judge (2001), se puede hablar de dos fases en la historia de las investigaciones sobre la personalidad en el trabajo. La primera empezaría en torno a los años 20 del siglo pasado, en los comienzos de la concepción moderna de personalidad y terminaría en la segunda mitad de los años 80. En esta etapa, predominaban los estudios primarios que intentaban relacionar diferentes inventarios de personalidad con distintos aspectos de la conducta humana, como algunos comportamientos organizacionales, entre los que destaca el desempeño laboral. Aunque la segunda guerra mundial intensificó el interés por la personalidad como predictor del desempeño, la conclusión general de los estudios de esta época fue que las medidas de personalidad no eran muy útiles para predecir el éxito laboral. Los autores de la primera revisión, Ghiselli y Barthol (1953), que abarcaba los estudios en los que las medidas de personalidad se utilizaron con fines de selección, recomendaron precaución a la hora de utilizarlos como predictores. Este recelo se debió al hallazgo de que los coeficientes variaban mucho entre diferentes profesiones. Revisiones posteriores llegaron a conclusiones aún menos optimistas. En este sentido,

Locke y Hulin (1962) concluyeron que la personalidad es un concepto sin ninguna utilidad práctica para la selección de personal. También Guion y Gottier (1965) se mostraron escépticos en cuanto a la utilidad de personalidad en la psicología organizacional aunque concluyeron que, en algunas situaciones y para algunos objetivos, algunos de los inventarios podrían resultar útiles. Resaltaron también la necesidad de llevar a cabo, en cada caso, estudios de las propiedades psicométricas de las medidas de personalidad antes de su aplicación. Sin embargo, cabe destacar que, a pesar de los recelos de los investigadores, los profesionales de recursos humanos siguieron utilizando los inventarios de personalidad durante todo este tiempo (véase Salgado, Ones y Viswesvaran, 2001), ajenos, y a veces indiferentes, al debate científico.

La falta de entusiasmo entre los investigadores de la psicología del trabajo respecto a las propiedades predictivas de las medidas de personalidad coincidió en el tiempo con una serie de controversias en el seno de psicología de personalidad que desafiaron la base de la evaluación de personalidad. Se trata de la llamada *revolución de Mischel* que llevó a muchos científicos a la conclusión de que la personalidad no se puede medir debido a que ésta *no existe.* Concretamente, Mischel (1968, 1973) afirmó que las acciones humanas no están determinadas por la personalidad de los individuos sino por los *factores situacionales* y éstos son fuerzas extrañas, indefinidas e invisibles.

Desde finales de los años sesenta y durante los siguientes veinticinco años, las conclusiones de los psicólogos organizacionales respecto al poder predictivo de la personalidad en el trabajo no cambiaron, prácticamente, (con algunas excepciones, p. ej., Elshout y Akkerman, 1975; R. T. Hogan, 1982, 1986). No había un sistema de clasificación que redujese los rasgos de personalidad hasta un número más manejable y, además, se utilizaban muchos rasgos distintos. No estaba del todo claro qué rasgos deberían medirse, así pues, se utilizaban distintos nombres para rasgos que medían lo mismo, y los mismos nombres para rasgos que medían constructos distintos (Guion y Gottier, 1965). Incluso, es más, cada rasgo era tratado más bien como un constructo independiente que como elemento de un factor superior. Tampoco se establecían las relaciones entre la personalidad y el desempeño basadas en un análisis del puesto. Las conclusiones no pudieron haber sido más desalentadoras.

Sin embargo, a mediados de los años 80 la visión de la personalidad entre los investigadores empezó a cambiar y las conclusiones de los nuevos estudios fueron un poco más optimistas. Este giro en el modo de percibir la personalidad dio comienzo a la segunda fase en su historia como predictor de comportamientos organizacionales, fase que dura hasta la actualidad. El momento más decisivo en la historia de la personalidad en el ámbito organizacional llegó, sin embargo, con el consenso de muchos investigadores sobre una nueva taxonomía de personalidad llamada el *Modelo de los Cinco Factores* (*Five Factor Model* – FFM) o bien, los *Cinco Grandes* (*Big Five* - BF). Este modelo presenta la personalidad como una estructura jerárquica con cinco factores de orden superior que abarcan los rasgos más específicos. Los Cinco Grandes fueron abrazados por gran parte de la comunidad científica, entre otros también por los psicólogos organizacionales, aunque no fueron desarrollados dentro de la psicología de organizaciones. Sin embargo, los investigadores se dieron cuenta de que solamente un modelo descriptivo de personalidad podría ayudar a estructurar los resultados provenientes de los diferentes autores y hacer que " ... *la torre Babel se convirtiera en una lengua común de toda comunidad* " (John y Srivastava, 1999, pág. 103).

A lo largo de los años 90 puede hablarse del *gran regreso* de la personalidad a la psicología de las organizaciones. Éste fue promovido por la mencionada aparición de los Cinco Grandes como propuesta de un modelo universal de personalidad y por la premisa de que las medidas de personalidad construidas de manera correcta predicen el desempeño laboral casi igual de bien que las medidas de habilidad cognitiva y que, además, no presentan el impacto adverso (R. T. Hogan, 2005). Adicionalmente, la aparición del nuevo modelo coincidió en el tiempo con la popularización del *meta-análisis* como un método de acumulación de datos a través de estudios (Hunter y Schmidt, 1990). Ambos acontecimientos fueron decisivos para que la personalidad pudiese recuperar su estatus perdido como un concepto relevante para la psicología en general y para la psicología de organizaciones en particular.

Principios de los Cinco Grandes

Enfoque psicoléxico

El modelo de los Cinco Grandes surgió del enfoque denominado *psicoléxico.* El pilar principal de este enfoque es la afirmación de que cada comportamiento significativo para la sociedad está incluido dentro lenguaje utilizado por sus miembros. Esta premisa se conoce como la *hipótesis léxica* en la literatura científica. Ésta fue definida por Goldberg en los siguientes términos:

> *Las diferencias individuales que son más significativas en las transacciones cotidianas entre las personas serán finalmente codificadas en su lenguaje. Cuanto más importante es esta diferencia, la gente la notará más y querrá hablar de ella, consecuentemente, será ideada una palabra para nombrarla*. (1981, págs. 141-142)

Desde los primeros estudios léxicos (Allport y Odbert, 1936; Baumgarten, 1933) hasta el relativo consenso sobre los Cinco Grandes (Digman, 1990; John, 1990) han pasado varias décadas de intensa investigación de los descriptores de personalidad enraizados en el lenguaje natural. Gran parte de esta primera investigación se llevó a cabo en inglés. La primera taxonomía en este idioma fue creada por Allport y Odbert en 1936. Dichos autores encontraron 17.953 términos en inglés referidos a la conducta humana. A continuación, clasificaron todas estas palabras en cuatro categorías: *rasgos personales*, *estados temporales*, *valoraciones sociales* y *términos metafóricos o dudosos*.

En un estudio posterior, Cattell (1943) utilizó la lista de 4.504 palabras incluidas en la primera de las cuatro categorías definidas por Allport y Odbert (*rasgos personales*), a la que añadió también algunos de los términos incluidos en la segunda categoría de *estados temporales*. Posteriormente, redujo esta lista a 150 grupos de rasgos, la mayoría presentados de forma bipolar. Además, se agregaron 21 categorías nuevas, fruto de la búsqueda bibliográfica. A continuación estas 171 fueron reducidas a 67. El objetivo de Cattell era poder llevar a cabo un análisis factorial en todo

momento, un método más objetivo que el simple juicio semántico. Cabe destacar que en aquellos tiempos todos los cálculos se hacían a mano, por lo que 67 *clusters* de rasgos era un conjunto todavía muy elevado para conseguir este fin. En un estudio posterior, aludiendo a la *" ... falta de medios para llevar a cabo un proyecto de tal alcance"*, (Cattell, 1945, pág. 71) decidió reducir todavía más el número de conjuntos de rasgos utilizados para el análisis final, refiriéndose a este proceso de reducción como *"la cuestión de triste necesidad"* (pág. 70). El resultado final del proceso de condensación fueron 35 grupos de rasgos o escalas que formaron un punto de partida para los análisis posteriores, mediante los que se encontraron 12 factores en una muestra de 208 varones adultos.

En los años posteriores, Cattell (1947, 1948) realizó una serie de análisis adicionales, en muestras tanto de varones como de mujeres, encontrando siempre 11 o 12 factores similares. Más tarde, añadió cuatro factores nuevos (Cattell, Eber y Tatsuoka, 1970), fruto de la búsqueda bibliográfica de los estudios dedicados a los cuestionarios de personalidad existentes. El conjunto de 16 factores, así creado, fue la base para la creación de un nuevo inventario de la personalidad llamado *Cuestionario de Personalidad de 16 Factores* (*16PF*). Sin embargo, ningún estudio independiente consiguió reproducir la estructura de 16 factores de Cattell (p. ej., Howarth, 1976; Howarth y Browne, 1971; Matthews, 1989), probablemente por ser el fruto de la mezcla entre el análisis del lenguaje natural y el análisis de los inventarios de personalidad. Finalmente, cabe resaltar que aunque muchos consideran a Cattell el "padre intelectual" del modelo de los Cinco Grandes (véase Goldberg, 1993), este siempre negó su relación con este modelo, manteniendo invariablemente que existían muchos más factores que cinco.

Continuando con el trabajo del análisis factorial de los grupos de rasgos encontrados por Cattell en 1945, Fiske (1949) utilizó 22 de las 35 escalas originales. Se basó para ello en una muestra de 128 estudiantes de psicología clínica. Obtuvo autovaloraciones, valoraciones de los compañeros y de psicólogos (observadores) para cada una de las 22 escalas y encontró una estructura persistente de cinco factores en los tres conjuntos de datos. A partir del estudio de Fiske, el principal problema de los seguidores del enfoque piscoléxico fue, en realidad, el número de factores que mejor

explicaban todo el dominio de la personalidad, ya que las variables de entrada casi siempre eran muy similares: Las 35 escalas de Cattell o bien un subconjunto elegido de las mismas (De Raad, 2000).

El hallazgo de los cinco factores convirtió a Fiske en el primer *descubridor accidental* del concepto de los Cinco Grandes, similar al que conocemos hoy en día (Goldberg, 1993). Aunque ya más de una década antes Thurstone (1934) sugirió con una sorprendente perspicacia que con cinco factores se puede describir adecuadamente el dominio de personalidad, al basarse en una lista de solo 60 adjetivos elegidos de manera algo aleatoria. Entonces, podría no tratarse del mismo modelo de cinco factores que el descubierto por Fiske (1949), después de una quincena de años de estudios sobre una taxonomía de los descriptores de personalidad en el lenguaje natural.

Una década después del descubrimiento de Fiske, los investigadores aún no estaban seguros de si cinco era el número de factores adecuado para describir de manera satisfactoria la personalidad. Para confirmar de manera definitiva el número óptimo de factores que abarcarían todo el dominio de personalidad, Tupes y Christal (1961) realizaron ocho análisis factoriales. En cada una de las ocho muestras empleadas encontraron cinco factores ortogonales y recurrentes, distintos entre sí pero muy similares entre las diferentes muestras. Estos cinco factores fueron denominados: Dominancia/Energía (*Surgency*) (1), Amigabilidad (2), Confiabilidad (*Dependability*) (3), Estabilidad Emocional (4) y Cultura (5).

Los estudios de Tupes y Christal no fueron tan conocidos como los estudios de Cattell, Fiske o posteriormente de Norman, debido a que habían sido publicados como informes técnicos de las fuerzas aéreas estadounidenses, siendo por lo tanto, accesibles a un público muy reducido. A pesar de esto, Goldberg (1993) "bautizó" a Tupes y Christal como los "verdaderos padres" (pág. 27) del modelo de los Cinco Grandes.

Una nueva etapa en los estudios psicoléxicos empezó con un estudio de Norman (1963) que investigó la razón de resultados diferentes en cuanto al número de factores obtenidos por Cattell (1947, 1948) y Tupes y Christal (1961). Norman eligió las cuatro variables que mostraron la mayor carga factorial en cada uno de los cinco

factores encontrados con anterioridad en un estudio de Tupes y Christal (1958), una versión reducida del estudio de 1961. De esta forma, Norman creó un subconjunto de 20 variables presentadas de forma bipolar (polos A y B). Estas variables fueron aplicadas en cuatro muestras de estudiantes universitarios (N=622). En todos los conjuntos de datos, Norman encontró la confirmación de la estructura de cinco factores relativamente ortogonales y fácilmente interpretables.

Las escalas utilizadas por Norman (1963) constituyen uno de los primeros instrumentos de evaluación de cinco factores, ampliamente utilizado por otros autores. Los factores encontrados (polos A) fueron denominados Dominancia/Energía (*Surgency*) (1), Amigabilidad (2), Consciencia (3), Estabilidad Emocional (4) y Cultura (5), y a partir de aquel momento, ganaron en la literatura la denominación de *los Cinco de Norman* (*Norman's Five*) (De Raad y Perugini, 2002) o bien *Cinco Grandes* (*Big Five*) (Goldberg, 1981). En cuanto a la primera denominación es necesario resaltar que -según el mismo Norman- estos factores deberían denominarse *los Cinco de Tupes y Christal* (*The Tupes and Christal Five*), debido a que él solo confirmó la estructura encontrada anteriormente por los mismos. En cuanto a la segunda denominación, es importante reconocer que estos factores fueron calificados como *grandes*, no para destacar su grandeza intrínseca, sino para resaltar su amplitud conceptual.

Norman (1963), sin embargo, se mostró escéptico con el hecho de que los factores encontrados representasen adecuadamente a todo el dominio de personalidad. Finalmente, constató que "*ya ha llegado la hora de volver al conjunto total (total pool) de las denominaciones de rasgos en el lenguaje natural*" (Norman, 1963, pág. 582) y decidió empezar de nuevo el trabajo de Allport y Odbert (1936) de creación de una taxonomía de rasgos de personalidad codificados en el lenguaje natural. En realidad, Norman esperaba encontrar un número de factores superior a cinco; no obstante, las investigaciones posteriores solo confirmaron cinco como el número óptimo de los factores de personalidad (p. ej., Borgatta, 1964a, 1964b; G. M. Smith, 1967, 1969).

En un trabajo posterior, Norman (1967) amplió el conjunto de 17.953 rasgos de Allport y Odbert (1936). Encontró en total casi 27.000 descriptores de personalidad, pero la mayoría resultaron ser simplemente variaciones de los rasgos incluidos en el

listado original de Allport y Odbert. Un largo proceso de categorización y de reducción de estos rasgos, prescindiendo de sinónimos, palabras ambiguas, desconocidas, etc., llevó a Norman a centrarse en 2.797 rasgos incluidos en la categoría de *rasgos estables*. A continuación, dichos rasgos se agruparon en 10 conjuntos, cada uno de los cuales indicaba alguno de los polos opuestos de *los Cinco de Norman* (Norman, 1963). Finalmente, se encontraron 75 grupos de sinónimos que abarcaban 1.431 adjetivos y 175 sustantivos que formaran la llamada *taxonomía de Norman*. Sin embargo, Norman ha dejado para otros investigadores el trabajo de refutar o bien de sostener la existencia de cinco grandes factores como un nuevo modelo de personalidad, ya que nunca continuó sus primeros análisis.

En este punto, cabe destacar el trabajo de Elshout y Akkerman (1975), dos psicólogos holandeses que durante su servicio militar tuvieron conocimiento mediante informes militares sobre los trabajos de Tupes y Christal (1958, 1961), desconocidos para el público general. Estos investigadores desarrollaron *Vijf Persoonlijkheids-Factoren Test*, 5PFT (*Five Personality-Factor Test, 5PFT*), el primer inventario de personalidad en el mundo creado, específicamente, para la medición de los Cinco Grandes, antes de que estos fueran denominados como tales, e incluso antes de que este modelo ganase reconocimiento mundial. Desde entonces, la comunidad científica holandesa y flamenca utiliza este cuestionario con éxito, siendo algunos ejemplos los estudios de Busato, Prins, Elshout y Hamaker (2000) o Klehe y Anderson (2007). El 5PFT se compone de 70 ítems, 14 para cada uno de los factores definidos como: Extraversión, Sociabilidad (Amigabilidad), Conciencia, Neuroticismo y Cultura. Además, en un estudio posterior se halló que estas grandes dimensiones abarcan en total 13 facetas o factores primarios (Elshout, 1999).

El siguiente gran personaje en la historia del desarrollo del nuevo modelo de personalidad, Goldberg, se mostró al principio igual de incrédulo que Tupes, Christal y Norman con el modelo de los Cinco Grandes (Goldberg, 1993), optando a su vez por un modelo de menos de cinco factores (en concreto por un modelo de tres factores). No obstante, finalmente llegó a ser un personaje de referencia en los estudios psicoléxicos y fiel partidario del modelo de los Cinco Grandes. Goldberg partió de la taxonomía desarrollada por Norman (1967) y, aunque llevó a cabo sus primeros estudios en los

años 80 (p. ej., Goldberg, 1981, 1982), no fue hasta el año 1990 cuando realizó su aportación más significativa para el desarrollo del nuevo modelo de personalidad. Su objetivo era construir su propio listado de rasgos, partiendo de la taxonomía de Norman (1967) y del Listado de Adjetivos de Gough (Gough y Heilbrun, 1965). El listado resultante de 1.710 rasgos abarcaba a todos los 1.431 adjetivos recogidos en la taxonomía de Norman (1967). Goldberg (1990) presentó su taxonomía de rasgos de forma apropiada para la autovaloración y recogió los datos de 187 sujetos para realizar tres estudios independientes. En el primero, los datos fueron segregados en 75 grupos (*clusters*) basados en los Cinco de Norman y sometidos a 10 tipos distintos de análisis factorial (utilizando cinco métodos diferentes de extracción de los factores y dos métodos de rotación: *varimax* y *oblimin*). Surgió la estructura de cinco factores generales en todos; incluso cuando se rotaron más de cinco factores, estos cinco primeros permanecieron invariables. Entre los factores con la misma denominación encontró altas correlaciones, independientemente del método utilizado para extraerlos y rotarlos (de .96 a .99). En el segundo estudio agrupó los 1.710 rasgos en 133 grupos de sinónimos y volvió a analizarlos utilizando los datos de autovaloraciones (dos muestras) y de valoraciones realizadas por los compañeros (dos muestras). Otra vez, el modelo de cinco factores principales surgió claramente. El tercer estudio consintió en determinar los rasgos que mejor representan (*markers*) estos cinco factores. Los factores fueron llamados: *Dominancia/Energía* (*Surgency*) (alegría, sociabilidad vs. distanciamiento, sosiego [*silence*]), *Amigabilidad* (cooperación, amabilidad vs. beligerancia, criticismo exagerado [*overcriticalness*]), *Conciencia* (organización, eficiencia vs. desorganización, negligencia), *Estabilidad Emocional* (placidez, independencia vs. inseguridad, miedo), y *Cultura (Intellect)* (intelectualidad, profundidad vs. superficialidad, falta de imaginación). Dos años más tarde, el modelo de cinco factores emergió también en los datos del nuevo estudio de Goldberg y sus colaboradores (Hofstee, De Raad y Goldberg, 1992), lo que aumentó todavía más el apoyo para el nuevo modelo de personalidad en la comunidad científica.

¿Son los Cinco Grandes universales?

En otros idiomas diferentes al inglés, los estudios psicoléxicos empezaron en los años 70, pero no fue hasta los años 80 cuando su desarrollo se intensificó. Se crearon taxonomías de adjetivos-descriptores de rasgos de personalidad en varias lenguas y se pretendía buscar la estructura factorial característica para cada idioma en particular. Los primeros estudios de este tipo se llevaron acabo en holandés (Brokken, 1978; De Raad, Hendriks y Hofstee, 1992; De Raad, Mulder, Kloosterman y Hofstee, 1988; Hofstee et al., 1992) y alemán (Angleitner, Ostendorf y John, 1990; Borkenau y Ostendorf, 1990; Ostendorf, 1990), idiomas muy próximos al inglés en cuanto a su génesis. Sin embargo, este fue solo el principio de las investigaciones que tiempo después se llevaron a cabo en otros idiomas, tales como húngaro (Szirmák y De Raad, 1994), italiano (Caprara y Perugini, 1994), checo (Hrebícková y Ostendorf, 1995), polaco (Szarota, 1995, 1996; Szarota, Ashton y Lee, 2007), ruso (Shmelyov y Pokhil'ko, 1993), español (García, Aluja y García, 2004; Quevedo-Aguado, Iraegui, Anivarro y Ross, 1996), francés (Boies, Lee, Ashton, Pascal y Nicol, 2001), portugués (Hutz et al., 1998), filipino (Church, Katigbak y Reyes, 1996, 1998; Church, Reyes, Katigbak y Grimm, 1997), chino (Yang y Bond, 1990), japonés (Aoki, 1971; Isaka, 1990), turco (Somer y Goldberg, 1999), hebreo (Almagor, Tellegen y Waller, 1995), griego (Saucier, Georgiades, Tsaousis y Goldberg, 2005), etcétera.

La mayoría de estos estudios acabaron confirmando la estructura de cinco grandes factores, como es el caso del holandés, alemán, húngaro, italiano, checo, polaco y filipino, aunque las interpretaciones de algunos factores varían levemente. La conclusión general de los estudios psicoléxicos internacionales es que, aunque resulta prácticamente imposible encontrar una estructura de rasgos idéntica en todos los idiomas, las estructuras encontradas son muy semejantes. Cuando el número de factores hallados era de cinco, el más problemático siempre era el quinto factor (De Raad y Van Heck, 1994; Di Blas y Forzi, 1998, 1999; Szirmák y De Raad, 1994). Este es el que más difiere entre idiomas; su contenido a veces es distinto en un idioma que en otro, al igual que varían sus denominaciones, por ejemplo, en alemán lleva nombre de

Intelecto, en holandés de Inconformidad (*Inconventionality*) y en italiano de Rebeldía. A veces, cuando el número de factores encontrados era superior a cinco, dos factores correspondían a uno de los Cinco Grandes, pero en ocasiones se necesitaban más de cinco factores para definir satisfactoriamente el dominio de personalidad (p. ej., Benet-Martínez y Waller, 1997; Church et al., 1997; Isaka, 1990; Katigbak, Church y Akamine, 1996). En general, la estructura más semejante se encontró entre las lenguas germánicas y más ampliamente, entre las lenguas europeas. Los resultados de los estudios de las lenguas no occidentales no son tan evidentes, ya que a veces menos o más de cinco factores definen mejor el dominio de personalidad. Sin embargo, las investigaciones han demostrado que la semejanza entre las estructuras encontradas no permite concluir que cada idioma o cultura refleje un modelo de personalidad diferente. Por lo tanto, aunque la conclusión sobre la universalidad lingüística de los cinco grandes factores derivados de los estudios psicoléxicos sería prematura, es la hipótesis que hasta el presente funciona mejor dentro del enfoque de rasgos (John, Naumann y Soto, 2008).

Enfoque disposicional

Al mismo tiempo que se desarrollaba el enfoque psicoléxico, algunos investigadores buscaban un marco común de la personalidad mediante el análisis factorial de los inventarios de personalidad existentes. El examen de los cuestionarios de personalidad existentes mediante el análisis factorial con el fin de encontrar el modelo de personalidad, ha sido denominado en la literatura como enfoque disposicional (Hrebícková et al., 2002) o bien, de manera más general, enfoque *etic* (J. W. Berry, 1969; Rolland, 2002). A su vez, la búsqueda del modelo de personalidad analizando el lenguaje natural se denominó enfoque léxico (Hrebícková et al., 2002) o, de manera más general, enfoque *emic* (J. W. Berry, 1969; Rolland, 2002).

Uno de los primeros investigadores en aplicar los métodos estadísticos en la búsqueda de los factores comunes entre varios inventarios de personalidad fue Eysenck. En 1947, éste sugirió que existían dos factores principales de personalidad:

neuroticismo, es decir, la tendencia a experimentar las emociones negativas, y extraversión, la tendencia a disfrutar de toda clase de eventos sociales. A pesar de la falta de acuerdo sobre el universalismo de estos factores, en la literatura científica apareció una denominación de este modelo como *Dos Grandes* (*Big Two*; Wiggins, 1968). A finales de los años 70, Eysenck constató que existía una tercera dimensión de personalidad que denominó *Psicoticismo* (Eysenck y Eysenck, 1977) y que sería una combinación, entre otros, de rasgos tales como la agresión, la impulsividad, el egocentrismo, la falta de empatía y la frialdad de emociones. El modelo de Eysenck en la literatura científica empezó a ser denominado PEN, de las iniciales de los tres factores que forman parte del mismo.

A pesar de los avances en esta línea de investigación, hasta los años 80 no se alcanzó un acuerdo sobre si era más adecuado el modelo de dos o de tres factores, o bien si algún otro modelo reflejaba más adecuadamente la estructura de personalidad en los inventarios de personalidad existentes. En medio de esta polémica, Costa y McCrae (1976), investigadores interesados en la evolución de la personalidad a lo largo de la vida humana, llevaron a cabo un análisis de datos provenientes del inventario 16PF (Cattell et al., 1970) en una muestra de voluntarios participantes de un estudio longitudinal de la Administración de Veteranos en Baltimore. Los resultados de este estudio les condujeron a la conclusión de que la personalidad se compone de tres factores. Dos de los factores encontrados eran muy similares al neuroticismo y extraversión de Eysenck (1947), por lo que los autores mantuvieron ambas denominaciones. El tercer factor fue denominado *apertura a la experiencia*, ya que contenía escalas tales como *imaginativo* y *aventurero*, entre otras.

Al inicio de los años 80, Costa y McCrae observaron la similitud de su modelo de tres factores con algunos factores del modelo de los Cinco Grandes que emergió del enfoque psicoléxico. La *extroversión* de la que ellos hablaban era claramente el factor de *dominancia/energía* de Norman (1963), al igual que *neuroticismo* era exactamente el lado opuesto de *estabilidad emocional*. Si el factor de *cultura* de Norman correspondiera al factor de *apertura a la experiencia*, su modelo sería un fiel reflejo de tres de los Cinco Grandes, prescindiendo únicamente de los factores: *conciencia* y *amigabilidad*. En una serie de estudios llevada a cabo por estos autores (McCrae y

Costa, 1985a, 1985b, 1987) se confirmó, finalmente, que *apertura a la experiencia* era muy similar a *cultura* de Norman. Este dato convenció a Costa y McCrae para añadir a su modelo los dos restantes factores y los convirtió en fieles partidarios del modelo de los Cinco Grandes.

La esencia de los Cinco Grandes

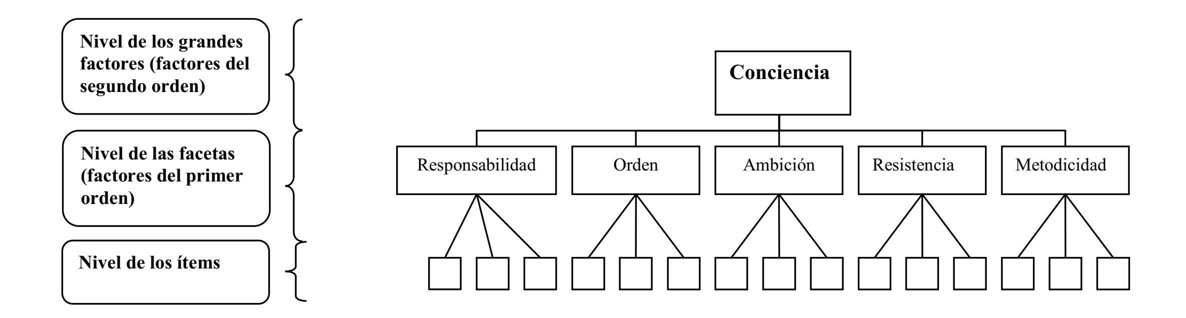
El modelo de los Cinco Grandes tal y como se conoce hoy en día supone la existencia de cinco factores ortogonales que son las fuentes más significativas de varianza existente entre múltiples descriptores de atributos de personalidad existentes en varios idiomas en todo el mundo (según el enfoque psicoléxico) o en diferentes inventarios de personalidad existentes (según el enfoque disposicional). Por factor se entiende aquí una reducción de múltiples variables observables a una variable latente, creada a partir de un conjunto concreto de datos (Saucier y Goldberg, 2002).

Dentro del modelo de personalidad de cinco factores existen dos puntos de partida para la descripción más detallada de su estructura: la perspectiva jerárquica, proveniente del enfoque disposicional, y la circular, proveniente del enfoque léxico. Aunque es cierto que ambas reconocen que los factores son dimensiones amplias de personalidad dentro de las cuales se recogen las subdimensiones más específicas (facetas), estas últimas son vistas de manera diferente. En el primer caso, las facetas se consideran jerárquicamente anidadas en los Cinco Grandes y, en el segundo caso, se ven como mezclas de los mismos.

En el caso de ambas perspectivas, en el nivel más bajo se encuentran los pensamientos, sentimientos y conductas reflejadas en los ítems específicos. Sin embargo, dentro de la perspectiva jerárquica, las respuestas a estos ítems se agrupan y convierten en indicadores de atributos específicos a los que se puede referir como subdimensiones o facetas de personalidad. Las facetas que comparten el mismo significado psicológico y, a menudo, la misma etimología (origen) se combinan para definir los grandes factores de personalidad (véase Figura 1). Aunque los investigadores no están de acuerdo sobre el número exacto de facetas que incluye cada factor, las facetas tienen en común que la mayor carga factorial siempre se da en

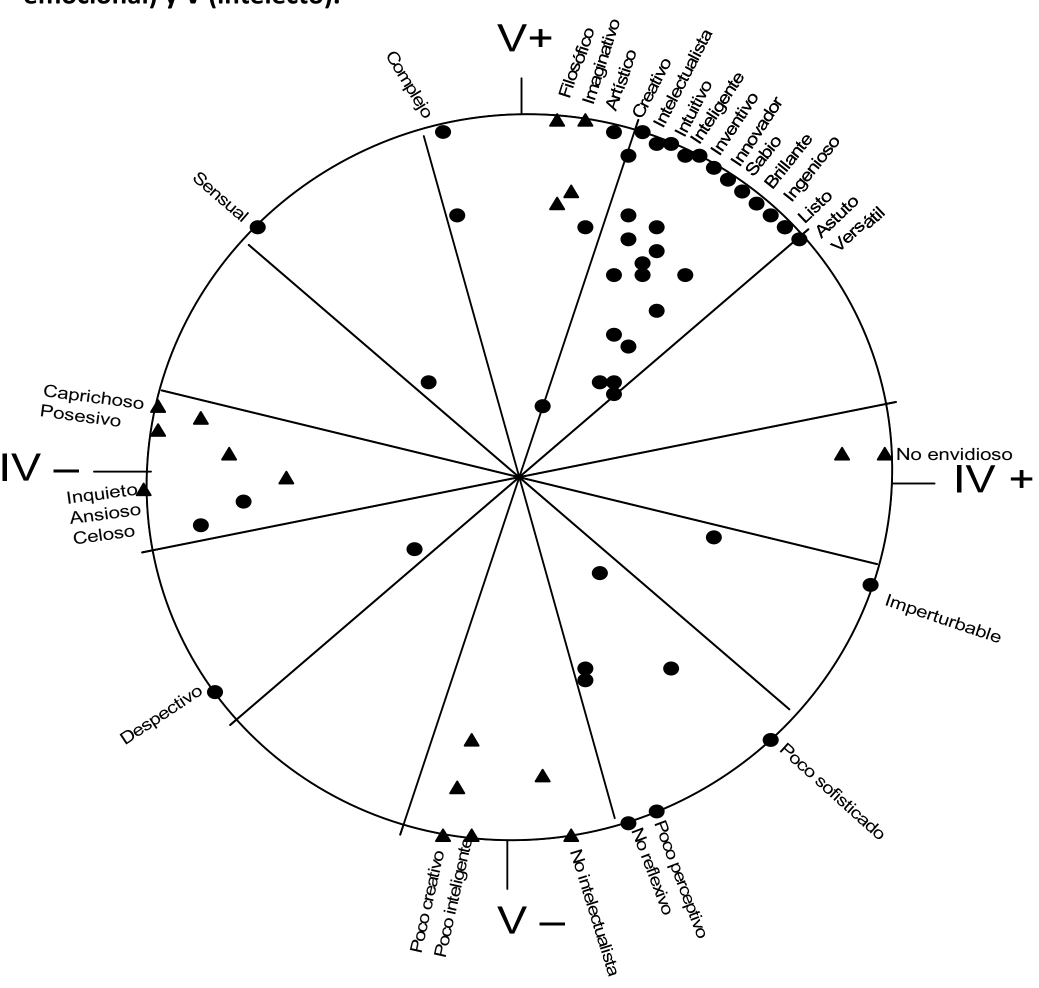
uno de los Cinco Grandes (llamado el factor primario). Además, aunque tuviesen cargas factoriales secundarias o terciarias en otros factores, la primera carga siempre es la más evidente.

Figura 1. Ejemplo de esquema de la estructura jerárquica del factor de conciencia.



Una representación alternativa a la estructura jerárquica de personalidad son los modelos circulares (p. ej., Peabody y Goldberg, 1989; Wiggins, 1982). En ellos, los rasgos específicos de personalidad se caracterizan por sus posiciones angulares en los espacios bidimensionales. El modelo circular más completo, a su vez basado en los Cinco Grandes factores, es el *Abridged Big Five Circumplex* (AB5C; Hofstee et al., 1992). Representa una estructura de facetas detallada, creando 10 espacios circulares para las posibles combinaciones de parejas de factores y diferenciando 90 segmentos unipolares (denominados facetas), que pueden convertirse en 45 segmentos bipolares. En el AB5C dominan los rasgos que constituyen una mezcla de dos factores, basándose en la observación de que la mayoría de rasgos tienen cargas factoriales significativas en dos y no en un solo factor. Cuando estos rasgos (en forma de adjetivos) se colocan en un espacio bidimensional utilizando sus cargas factoriales como coordinadas (puntos de referencia), forman un círculo (véase Figura 2). Cada rasgo puede tener cargas factoriales en un solo polo de los factores (positivo o negativo) o bien tener carga positiva en uno de los factores y negativa en otro. De este modo, dentro de este modelo, cada rasgo está definido por sus dos cargas factoriales más altas.

Figura 2. Modelo circular de las facetas mezclas de factores IV (estabilidad emocional) y V (intelecto).



Notas: Los triángulos representan las localizaciones de los términos factorialmente unívocos cuyas cargas secundarias son mínimas.

Fuente: Hofstee y colaboradores (1992).

Sin embargo, el AB5C se ha encontrado con la crítica de los investigadores provenientes del enfoque disposicional. Costa y McCrae (1995) argumentaron que este modelo no cumple con el supuesto de que las facetas deben tener una amplitud comparable; muchos de los segmentos contienen únicamente unos pocos rasgos, mientras que otros son muy amplios. Cabe destacar que existen 43 rasgos que

combinan baja amigabilidad con alto neuroticismo, pero solamente 13 que combinan baja amigabilidad como bajo neuroticismo o alta amigabilidad con alto neuroticismo. Además, este sistema no consigue hacer distinciones entre los rasgos que, a pesar de tener la misma localización en el modelo circular, son substancialmente diferentes. Por ejemplo, las escalas de gregarismo y emociones positivas del NEO PI-R son descriptores relativamente puros de extraversión y, como tales, estarían asignadas al mismo segmento por el modelo circular. Sin embargo, ambos rasgos se refieren a características totalmente distintas (preferencia por las interacciones interpersonales frente al nivel de alegría característico de una persona y, en realidad, están solo ligeramente correlacionadas (.35) (Costa y McCrae, 1992b). Pese a esta crítica, los seguidores del enfoque disposicional no han ofrecido una alternativa unánime al modelo circular, definiendo un número diferente de rasgos específicos y denominándolos de forma distinta en cada caso. Por todo lo anterior, una estructura de personalidad uniforme en el nivel de facetas se convirtió en uno de los objetivos más demandados por la comunidad científica hoy en día (véase Capítulo 2 del presente trabajo).

Aunque las diferencias entre los enfoques psicoléxico y disposicional han sido frecuentemente acentuadas por sus seguidores, sus frutos son tan similares que en el mundo científico se habla de un único modelo de personalidad. Las denominaciones provenientes de ambos enfoques, los Cinco Grandes (BF), el fruto del enfoque psicoléxico y el Modelo de Cinco Factores (FFM), que resultó del enfoque disposicional, se usan prácticamente como sinónimos. Además, los instrumentos que miden los cinco grandes factores provenientes de ambos enfoques muestran una sustancial validez convergente, por lo que existen pruebas no solo teóricas sino también empíricas que demuestran que, a pesar de utilizar vías distintas, las perspectivas psicoléxica y disposicional llegaron al mismo o muy semejante resultado final. A continuación, se presenta una breve definición de cada factor.

Definición de los factores

Estabilidad Emocional/Neuroticismo

A este factor le fueron asignados nombres distintos, por ejemplo "control emocional" (Fiske, 1949), *emocionalidad* (*emotionality*) (Borgatta, 1964a), *ansiedad* (Wiggins, 1968), *fuerza del ego* (*ego strenght*) (Digman y Takemoto-Chock, 1981) o *inestabilidad emocional* (De Raad et al., 1988), sin embargo, las dos denominaciones más comunes son *neuroticismo* y *estabilidad emocional* (que sería el lado opuesto de neuroticismo). A las personas que puntúan alto en estabilidad emocional (o bajo en neuroticismo) se las describe como pacientes, relajadas, poco excitables, imperturbables, no envidiosas y desapasionadas. A su vez, las que puntúan bajo en estabilidad emocional (o alto en neuroticismo) se las describe, generalmente, como emotivas, temperamentales, irritables, celosas, posesivas y de humor variable.

Las personas bajas en estabilidad emocional (altas en neuroticismo) se las describe también como más propensas a experimentar los sentimientos negativos, tales como temor, rabia, culpabilidad, nostalgia, vergüenza, disgusto. Estas emociones afectan de manera negativa a las capacidades de adaptación del individuo y están vinculadas con una menor capacidad de controlar sus impulsos. En comparación con la población media, las personas poco estables emocionalmente presentan mayores dificultades para afrontar las situaciones de estrés. A su vez, los que puntúan alto en estabilidad emocional (bajo en neuroticismo) suelen estar serenos y relajados, no suelen alterarse ni conmoverse en situaciones que pueden provocar una cierta tensión.

Extraversión

Este factor implica sociabilidad y facilidad para relacionarse con los demás, al igual que la preferencia por grupos y reuniones. A los extravertidos se les describe como personas activas y habladoras. Les gusta la excitación, la estimulación y tienden

a ser de carácter alegre. Se las caracteriza como personas animosas, enérgicas y optimistas. En el polo opuesto de la extraversión se encuentra la introversión. A las personas introvertidas se las describe más bien como reservadas; que no necesariamente tienen que sufrir ansiedad social, sino que a veces simplemente prefieren estar solas.

Dependiendo de la faceta en la que se había puesto el énfasis, los científicos denominaron este factor también como *autoexpresión segura* (*confident self-expression*) (Fiske, 1949), *dominancia/energía* (*surgency*) (Tupes y Christal, 1961), *dominancia/energía* o *extraversión* (Norman, 1963) o *dominancia/energía* como mezcla de potencia y actividad (Goldberg, 1981).

Apertura a la experiencia/Cultura

Este factor se denomina *apertura a la experiencia* dentro del enfoque disposicional. A su vez, dentro de la tradición psicoléxica la designación más común hoy en día es *cultura* (Norman, 1963; Tupes y Christal, 1961), aunque se han sugerido también distintas denominaciones del mismo, tales como: *inteligencia* (Borgatta, 1964a), *imaginación* (Saucier, 1994) o *autonomía* (Hendriks, Hofstee y De Raad, 1999).

Las personas que puntúan alto en apertura (cultura) se las describe como de fácil adaptación, que tienen apreciación por el arte, que muestran autorreflexión y preferencia por la diversidad y que se caracterizan por búsqueda intelectual. Las personas abiertas toman en consideración nuevas ideas y valores no convencionales, además de experimentar tanto las emociones positivas como las negativas de manera más intensa que las personas más cerradas. Además, se las describe como personas poco convencionales, propensas a cuestionar la autoridad y a aceptar fácilmente nuevas ideas éticas, sociales y políticas; sin embargo, no puede decirse que carezcan de principios. Por otro lado, las personas con puntuaciones bajas en la apertura a la experiencia (cultura) suelen describirse como personas que tienden a obedecer las normas, prefieren lo familiar a lo novedoso y, en comparación con las personas abiertas, con un abanico de intereses más reducido.

Amigabilidad

Este factor a lo largo de tiempo ha recibido denominaciones alternativas, tales como: *conformidad* (Fiske, 1949), *amable obediencia* (*freindly compliance*) frente a *hostil desobediencia* (*hostil noncompliance*) (Digman y Takemoto-Chock, 1981), o *amigabilidad* (*likeability*) (R. T. Hogan, 1986; R. T. Hogan y Hogan, 1992). Hoy en día, sin embargo, *amigabilidad* es la denominación más utilizada.

El núcleo de este factor lo forma el mantenimiento de las relaciones interpersonales, es decir, abarca los rasgos del individuo que le ayudan a convivir en comunidad, por ejemplo: ser compasivo, simpático, cariñoso, servicial, cooperativo y considerado. La persona amigable suele describirse básicamente como altruista, es decir, se entiende con los demás, está dispuesta a ayudarles y cree en que otras personas serán reciprocas. Por el contrario, la persona que puntúa bajo en este factor es desagradable, egocéntrica y suspicaz respecto a las intenciones de los demás, siendo reacia a cooperar.

Aunque el polo positivo de este factor parece socialmente más deseable y psicológicamente más saludable, ya que las personas amables suelen ser más populares que las personas antipáticas, la disposición para luchar por los propios intereses, a menudo, resulta ventajosa. Por ejemplo, la amigabilidad no siempre es una virtud en el campo de batalla o en un tribunal de justicia. A su vez, la actitud escéptica y crítica puede favorecer el avance científico. Por lo tanto, se puede constatar que intrínsecamente ninguno de los polos de este factor es el mejor.

Conciencia

El factor conciencia representa la fuerza para llevar a cabo las tareas y abarca aquellas características del individuo que son necesarias para conseguirlo, por ejemplo: ser organizado, sistemático, eficiente, práctico y constante (Goldberg, 1992).

Las personas que puntúan alto en el factor de conciencia suelen describirse como responsables, persistentes y decididas. Además, se consideran escrupulosas,

puntuales y fiables. Las que obtienen puntuaciones bajas, sin embargo, no es que ignoren las normas y los principios morales, sino que se las considera menos firmes en aplicarlos y más descuidadas a la hora de luchar para alcanzar sus metas. En general, altas puntuaciones en conciencia se asocian con un alto rendimiento académico (Noftle y Robins, 2007; Paunonen, 2003) o profesional (Barrick et al., 2001; Salgado, 2003c), pero cabe destacar que en su extremo pueden conducir a un extremado sentido crítico, a una escrupulosidad compulsiva o a una conducta de adicción al trabajo (Costa y McCrae, 1992b).

Debate alrededor del modelo de los Cinco Grandes

A pesar de que el modelo de los cinco grandes factores puede considerarse hoy en día el modelo de personalidad sobre el que existe mayor consenso entre los investigadores (Barrick y Mount, 2003; Soto y John, 2009), este desató también cierto debate. La mayor parte del mismo tuvo lugar a principios de los años 90, cuando los Cinco Grandes se encontraban todavía *en desarrollo*. Esto es, cuando se acumulaban las primeras evidencias acerca de su validez y de su generalización cultural. Sin embargo, en la actualidad este debate ha perdido peso, a pesar de que algunas de las cuestiones planteadas en aquel momento sigan abiertas y de que hayan surgido nuevas cuestiones de menor envergadura.

Al principio se criticó el origen empírico de los Cinco Grandes y la falta de soporte teórico, al igual que se puso en entredicho la razón de la hipótesis léxica. La variedad de las denominaciones de las grandes dimensiones también suscitó debate. Además, algunos investigadores no reconocen que sean cinco el número de factores definitivo, otros a su vez reprochan que los factores son demasiado generales para ser útiles en la predicción de los criterios organizacionales de interés. Por otro lado, las críticas principales en el ámbito de la Psicología del Trabajo y las Organizaciones, o de la selección de personal más particularmente, se centran hoy en día en los valores moderados de los coeficientes de validez encontrados en las investigaciones meta-analíticas.

McAdams (1992) resaltó que el surgimiento de los cinco grandes factores no fue apoyado en ninguna teoría. Consideró que únicamente, a posteriori, se han hecho

algunos esfuerzos para demostrar que los cinco factores que originalmente se derivaron del lenguaje natural sin ninguna base teórica, pueden ser encontrados en los inventarios de personalidad que sí la tienen. Además, reprochó que este modelo no posea una base biológica. Según él, lo que proporciona la legitimidad máxima a una taxonomía y lo que permite a sus constructos aportar las explicaciones causales es, precisamente la base biológica. Sin embargo, posteriormente, se encontró cierta evidencia de heredabilidad de los Cinco Grandes (Jang, Livesley y Vernon, 1996; Riemann, Angleitner y Strelau, 1997). Además, las investigaciones confirmaron que utilizando los Cinco Grandes las diferencias en las puntuaciones factoriales entre los miembros de culturas diferentes (Hough, Oswald y Ployhart, 2001), y dentro de culturas diferentes, entre ambos sexos (Costa, Terracciano y McCrae, 2001) son mínimas, también cuando los tests de personalidad se usan con los fines de selección de personal (Hough et al., 2001; Ones y Anderson, 2002).

En cuanto a la estabilidad del modelo a lo largo de la vida de las personas, hay dos investigaciones recientes que tratan de aclarar esta cuestión y que llegaron a conclusiones aparentemente opuestas. Terracciano, Costa y McCrae (2006) constataron en un estudio longitudinal que la personalidad después de los treinta permanece casi invariable (la relación promedio encontrada entre los puntos de medición ha sido .70, es decir muy alta). A su vez, Roberts, Walton y Viechtbauer (2006a) en un estudio meta-analítico hallaron que la personalidad puede cambiar en la vida adulta. Sin embargo, los tamaños de efecto (*d*) encontrados, a pesar de ser interpretadas como importantes, fueron en realidad muy pequeñas (varían de .06 a .30). Una diferente interpretación de los datos de ambos trabajos, desató un debate entre los autores (Costa y McCrae, 2006; Roberts, Walton y Viechtbauer, 2006b). En cambio, un análisis pormenorizado de sus resultados nos puede llevar a la reflexión de que la magnitud de los cambios encontrados en uno y otro estudio es, en realidad, similar. Los tamaños de efecto encontrados por Roberts y colaboradores (2006a) implican las diferencias de entre .06 y .30 desviaciones típicas entre las puntuaciones obtenidas por las personas de diferentes grupos de edad. Estas diferencias son tan pequeñas que puede concluirse que ambos estudios demuestran que la personalidad parece prácticamente estable a lo largo de vida adulta.

El modelo de los Cinco Grandes también permanece relativamente estable cuando se usan distintas fuentes de evaluación: autoevaluaciones, evaluaciones realizadas por los compañeros, conocidos, amigos y por los cónyuges. La relación entre ellas varía de .30 a .65, por lo que no puede ser fruto de la casualidad (Costa y McCrae, 1992b; Funder, Kolar y Blackman, 1995; Riemann et al., 1997). Finalmente, la estabilidad de este modelo se probó reproduciendo su estructura de manera satisfactoria en muestras de edades distintas y características diferentes (niños, adolescentes, universitarios, trabajadores adultos y personas mayores) (De Fruyt et al., 2006; De Fruyt et al., 2009; De Fruyt, Mervielde, Hoekstra y Rolland, 2000; McCrae et al., 2002; McCrae, Terracciano y 78 *members of the Personality Profiles of Cultures Project*, 2005; Mervielde y De Fruyt, 2000).

En cuanto a la hipótesis léxica, los críticos desconfiaban si realmente los factores provenientes del análisis del lenguaje natural son transferibles a otras culturas que no comparten el mismo origen lingüístico o la misma perspectiva social (Pervin, 1994). Se acentuó que la correspondencia de los Cinco Grandes en los idiomas no occidentales no es absoluta. Sin embargo, aunque muchos autores han encontrado más de cinco factores, ninguno de ellos ha sido reproducido a través de los distintos conjuntos de descriptores de personalidad (De Raad, 2000). Por todo lo anterior, aunque no se puede hablar de universalidad lingüística de los cinco grandes factores derivados de los estudios psicoléxicos, no hay una hipótesis alternativa que ganase más reconocimiento que los Cinco Grandes dentro del enfoque de rasgos (John et al., 2008).

La crítica anterior está directamente relacionada con el debate respecto al número definitivo de los factores en la jerarquía de personalidad. Si estos son cinco, los opositores querían saber: ¿Por qué cinco y no más o menos? (Block, 1995, 2001; McAdams, 1992). Insistían en que no existía ninguna prueba que justificara que cinco factores fuese un número adecuado para describir la personalidad humana. Sugerían que los análisis posteriores a Tupes y Christal (1961) moldearon de manera inconciente las replicas consecutivas de la estructura de los Cinco Grandes, debido a un previo compromiso con un número y una estructura particular. En muchos estudios se utilizó

el mismo conjunto de variables (p. ej., Cattell) y no era de extrañar que, basándose en la misma muestra, apareciera un modelo muy similar.

Como respuesta a lo anterior, tampoco se encuentran indicios de que cinco no sea un número adecuado de factores para describir la personalidad humana. Es más, desde estos primeros interrogantes se han llevado a cabo cientos de investigaciones en los conjuntos de variables totalmente independientes del conjunto original de Cattell, alcanzando siempre el mismo resultado. El modelo de cinco grandes factores ha sido corroborado de manera satisfactoria en múltiples idiomas, incluso no indoeuropeos, tanto desde la perspectiva del enfoque psicoléxico (p. ej., Almagor et al., 1995; De Raad, Perugini, Hrebícková y Szarota, 1998; García et al., 2004; Saucier et al., 2005; Somer y Goldberg, 1999), como desde el enfoque disposicional (McCrae et al., 2005).

Algunos investigadores, sin embargo, dirigieron la atención de la cuestión del número de factores a los diferentes niveles de abstracción en la jerarquía de los rasgos. Markon, Krueger y Watson (2005) consideran que en cada nivel de esta jerarquía existe un número diferente de rasgos y que todos los modelos alternativos a los Cinco Grandes pueden ser incorporados dentro de un único marco estructurado de forma jerárquica. Dentro de este marco existen factores más abstractos que los Cinco Grandes por encima de los mismos y más específicos por abajo. Esta perspectiva contradice las opiniones anteriores que resaltan que los cinco grandes factores son demasiado heterogéneos, por lo que oscurecen las relaciones entre la personalidad y los criterios de interés (Hough, 1998; Hough y Schneider, 1996; Paunonen y Jackson, 2000). Defendiendo la exhaustividad del modelo, John y colaboradores (2008) surgieren que los Cinco Grandes representan más bien un nivel alto en la jerarquía de la personalidad (aunque existan todavía los niveles superiores de la medición). Explican poniendo como ejemplo que son para la personalidad lo que las categorías de plantas y animales son para la biología: extremadamente útiles para algunas distinciones iniciales y aproximadas, aunque menos específicas para predecir las conductas puntuales.

En general, incluso los defensores de los Cinco Grandes están de acuerdo en que una taxonomía de rasgos completa debería incluir también las categorías de rasgos del nivel intermedio (facetas). Por ejemplo, Norman (1967) y luego Goldberg

(1982, 1990) desarrollaron entre 45 y 75 categorías subordinadas a las cinco grandes dimensiones. No obstante, estas categorías no fueron investigadas de manera sistemática, ni tampoco han sido incluidas en ningún instrumento evaluativo. Las facetas de NEO PI-R de Costa y McCrae (1985; 1989), o los compuestos de ítems homogéneos en el HPI (*Hogan Personality Inventory*) de R. T. Hogan (1986) y R. T. Hogan y Hogan (1992), han sido consideradas como pasos en la dirección correcta, aunque la asignación de los mismos a cada factor ha sido de algún modo arbitraria. Además, tanto el NEO PI-R (Costa y McCrae, 1985; 1989) o el HPI (1986; R. T. Hogan y Hogan, 1992) como otros instrumentos basados en los cinco grandes factores, difieren significativamente respecto al número y tipo de los rasgos intermedios (facetas). En este punto corresponde evocar las palabras de Block (1995, pág. 209): "*Tristemente para el progreso en la evaluación de personalidad, una vez que las 'dimensiones' de personalidad se profundizan y se ofrecen las subdimensiones o facetas, aparece toda clase de complicaciones, desorden y desacuerdos*". Esta cuestión indica, sin lugar a dudas, una necesidad del trabajo conceptual y empírico adicional, para llegar a un consenso en cuanto a la estructura de los cinco factores en el menor nivel de abstracción, una cuestión tratada con más detalle en el capítulo siguiente.

También las denominaciones de los Cinco Grandes desataron la polémica. Al principio, los opositores proclamaban que existían diferentes *Cinco Grandes* debido a que numerosos investigadores han utilizado distintas etiquetas y/o instrumentos, incluyendo Fiske (1949), Tupes y Christal (1961), Norman (1963), Borgatta (1964a), Digman y Inouye (1986), Costa y McCrae (1985), Conley (1985), De Raad y colaboradores (1988), Botwin y Buss (1989), Field y Millsap (1991), Trapnel y Wiggins (1990), y Goldberg (1990), entre otros. Sin embargo, John (1990) y Costa y McCrae (1992a) destacaron que, a pesar de las diferencias entre las diferentes denominaciones, los investigadores están de acuerdo en que se refieren al mismo fenómeno. Existe también una amplia evidencia empírica de que los instrumentos que miden este modelo demuestran una validez convergente (De Raad y Perugini, 2002).

Con el tiempo, la gran mayoría de los investigadores adaptaron las denominaciones provenientes de las primeras letras de los nombres de los factores propuestos por McCrae y Costa (1985b), por lo que las discrepancias antes

mencionadas desaparecieron. Además, John y colaboradores (2008) argumentan que sería muy difícil encontrar denominaciones que reflejasen mejor la esencia de cada factor. Resaltan que la principal ventaja de estas denominaciones es justamente su popularidad y uso frecuente, que permite que exista un marco común para todas las investigaciones sobre este modelo y que sean posibles las comparaciones entre los resultados obtenidos.

Finalmente, cabe destacar que recientemente las críticas se centran más que en el modelo de los Cinco Grandes en sí, en su utilidad en un contexto determinado, como, por ejemplo, en los procesos de selección de personal. Algunos investigadores resaltan que su utilidad es limitada, debido a que los índices de validez predictiva alcanzados por las medidas de personalidad son, en el mejor de los casos, moderados (Morgeson et al., 2007a, 2007b). Sin embargo, admitiendo su moderada magnitud, Ones, Dilchert, Viswesvaran y Judge (2007) argumentan que estos índices pueden ser relevantes, debido a que la personalidad presenta validez añadida a la mayoría de las pruebas de selección más comunes.

En conclusión, puede decirse que a pesar de que no existe unanimidad al respecto, este modelo está ampliamente compartido por la comunidad científica. Además, es el modelo de personalidad más exhaustivo y económico que existe hoy en día, al mismo tiempo que la evidencia de su utilidad es más voluminosa y convincente que para cualquier otro modelo (Dilchert, Ones, Van Rooy y Viswesvaran, 2006). Los Cinco Grandes son un marco muy valioso para resumir los resultados de las investigaciones realizadas en diversos países, en personas de diferentes sexos y en diferentes periodos de tiempo. Si juzgamos este modelo por los resultados que ofrece y no por la crítica que desató, no hay lugar a dudas de que la aparición de los *Cinco Grandes* fue algo beneficioso para toda la comunidad científica y para la psicología del trabajo en particular. Para corroborar lo antedicho, a continuación se detallan las aportaciones de este modelo para el entendimiento del papel que tiene la personalidad en las organizaciones.

Los resultados de la investigación sobre la personalidad en el trabajo

El interés de los psicólogos organizacionales por la personalidad proviene de la premisa de que las características humanas influyen en el desempeño laboral y en otros criterios relacionados con el trabajo. Sin embargo, antes de la aparición del modelo de los *Cinco Grandes*, el concepto de personalidad parecía no cumplir estas premisas. Las revisiones de la investigación realizadas en los años 50 y 60 del siglo XX causaron gran escepticismo entre los investigadores respecto al uso de personalidad en los contextos organizacionales (p. ej., Ghiselli y Barthol, 1953; Guion y Gottier, 1965; Locke y Hulin, 1962).

El estudio cuantitativo de Schmitt, Googing, Noe y Kirsch (1984) sobre las propiedades predictivas de varios tipos de predictores en el campo de psicología organizacional en cuanto a la personalidad, llegó a conclusiones similares a las revisiones cualitativas de literatura antes mencionadas. Los inventarios de personalidad alcanzaron un coeficiente de validez de .15 promedio (.20 para predecir las valoraciones de los superiores, .12 para predecir la rotación, .15 para predecir el logro/las notas, .12 para predecir el cambio de estatus y .27 para predecir el salario), el valor más bajo de todos los predictores analizados. Sin embargo, cabe destacar que Schmitt y colaboradores no utilizaron en su estudio ni la corrección por la restricción del rango, ni la corrección por la fiabilidad del criterio, además de que agruparon los coeficientes de validez sin tener en cuenta cuál ha sido realmente el constructo medido por los diferentes inventarios de personalidad. Por lo tanto, puede concluirse que los resultados de este análisis cuantitativo no reflejan las capacidades predictivas de las medidas de personalidad, subestimando la validez de las mismas.

El siguiente trabajo -cuyo objetivo fue verificar las propiedades predictivas de la personalidad- es el meta-análisis de Hough, Eaton, Dunnette, Kamp y McCloy (1990). Los autores basaron su taxonomía de las escalas de personalidad en la clasificación de seis factores de R. T. Hogan (1982). Dichos factores fueron denominados *dominancia/energía* (*surgency*), que corresponde al factor de *extraversión*; *adaptación* (*adjustment*), que corresponde al factor de *estabilidad emocional*; *amigabilidad* (agreeablness); *confiabilidad* (*dependability*), que corresponde al factor de *conciencia*;

intelectualidad (*intellectance*), que corresponde al factor de *apertura a la experiencia*; y *afiliación* (*afiliation*), que no está representada en el modelo de los *Cinco Grandes*. 146 escalas provenientes de los 12 inventarios de personalidad más comúnmente utilizados en aquel entonces fueron clasificadas según esta taxonomía de seis factores. Sin embargo, solo 117 escalas de las 146 escalas originales pudieron ser asignadas a alguna de las seis grandes dimensiones; las 29 escalas restantes fueron asignadas a la categoría adicional denominada *heterogénea* (*miscellaneous*). Entre las que formaron parte de los factores *dominancia/energía, adaptación, amigabilidad, fiabilidad, intelectualidad, afiliación* y la escala *heterogénea* se obtuvieron las correlaciones de .46, .43, .37, .34, .40, .33 y .05, respectivamente. Posteriormente, se dividió la categoría heterogénea en tres constructos independientes: *logro* (*achievement*), *masculinidad* (*masculinity*) y *locus de control* (*locus of control*). A continuación, se analizó la literatura científica desde el año 1960 a 1984 en búsqueda de estudios que reportasen coeficientes de correlación con los siguientes criterios: educación, formación, implicación en el trabajo (*job involvement*), desempeño laboral (*job proficiency*), delincuencia y abuso de sustancias nocivas para la salud (p. ej., drogas, alcohol, tabaco). Los cinco últimos criterios son comúnmente utilizados en psicología organizacional, siendo la delincuencia y el abuso de sustancias nocivas para la salud tipos de conductas contraproductivas.

Los resultados de este meta-análisis indicaron que adaptación (estabilidad emocional) y confiabilidad (conciencia) son buenos predictores del desempeño laboral (.13). Los mismos factores resultaron predecir la implicación en el trabajo, .13 y .17 respectivamente, además de Intelectualidad (apertura a la experiencia) (-.10). Para el criterio de formación se encontraron correlaciones de .16 con adaptación (estabilidad emocional), .14 con intelectualidad (apertura a la experiencia), .11 con confiabilidad (conciencia), .10 con amigabilidad, y .08 con dominancia/energía (extraversión). Las correlaciones más altas sin embargo correspondían al criterio de delincuencia. Hough y colaboradores (1990) encontraron que este criterio se relaciona con adaptación (estabilidad emocional) (-.43), amigabilidad (-.31), confiabilidad (conciencia) (-.27) e intelectualidad (apertura a la experiencia) (-.24). En el caso de abuso de sustancias, las relaciones moderadas se encontraron con los factores de confiabilidad (conciencia) (-.28) e intelectualidad (apertura a la experiencia) (.18). Resumiendo, dos factores,

adaptación (estabilidad emocional) y confiabilidad (conciencia) resultaron ser buenos predictores de todos los criterios organizacionales, aunque en el caso de adaptación la relación con el criterio de abuso de sustancias fue muy pequeña (-.07). Salgado (1997) destaca, pues, que los valores presentados por Hough y colaboradores no han sido corregidos por falta de fiabilidad en el criterio y restricción en el rango. Utilizando el valor de la fiabilidad del criterio de .52 (SD = .05) y la restricción del rango de .94 (SD = .05) utilizados en un meta-análisis posterior de Barrick y Mount (1991), la validez corregida de adaptación y confiabilidad para predecir el desempeño laboral sería .23, mientras que para el criterio de formación los coeficientes de validez corregidos serían de .28, .25, .19, .18 y .14 para adaptación, intelectualidad, fiabilidad, amigabilidad y dominancia/energía, respectivamente. Además, Hough y colaboradores no han indicado los tamaños de las muestras utilizados para alcanzar los coeficientes de validez predictiva arriba mencionados, lo que implica que no se puede llevar a cabo la comparación de este estudio con los meta-análisis posteriores.

Sin embargo, la visión de personalidad como predictor ineficaz del desempeño laboral no cambió hasta que fueron aplicadas las innovaciones metodológicas en el meta-análisis (principalmente correcciones por falta de fiabilidad en el predictor y en el criterio y por la restricción del rango) para los estudios basados en las medidas de personalidad. Esto fue posible gracias a la creciente utilización del modelo de los *Cinco Grandes* para clasificar los datos. En 1991, Barrick y Mount meta-analizaron 162 muestras provenientes de 117 estudios. La muestra total abarcaba 23.944 sujetos y fue dividida en los siguientes grupos ocupacionales: especialistas, policías, gerentes, vendedores, trabajadores cualificados y semi-cualificados. Se utilizaron tres criterios organizacionales: eficiencia en el trabajo, éxito en la formación y datos de personal (productividad, salario, permanencia en el puesto, etc.). Las escalas de los inventarios de personalidad empleados en los estudios analizados fueron clasificadas según el modelo de los *Cinco Grandes* por seis jueces, basándose en las descripciones de los factores presentadas por Digman (1990). Los jueces ubicaron cada escala dentro del factor que, según ellos, mejor le correspondía. Cabe destacar, por otra parte, que esta clasificación tenía ciertos problemas, ya que escalas que miden solo una parte del constructo fueron agregadas y tratadas como las medidas, con la validez de constructo completa. También podían ocurrir errores en la clasificación de dichas escalas. Estos

posibles errores artifactuales pueden causar una subestimación de la validez de los *Cinco Grandes*. Para evitar el primer problema sería necesaria la aplicación de la corrección de la validez por la medición imperfecta del constructo, sin embargo, ésta no se aplicó en este caso, aunque Barrick, Mount y sus colaboradores la aplicaron en su meta-análisis posterior (Mount y Barrick, 1995). No obstante, en el caso de incorrecta clasificación de escalas, no hay una corrección posible.

Barrick y Mount (1991) encontraron que el factor conciencia era un predictor válido en todos los grupos ocupacionales con una validez predictiva que varía de .20 (teórica) para profesionales a .23 para vendedores, aunque en el caso de policías el resultado no es generalizable ya que el intervalo de confianza incluía el cero. El resto de los factores alcanzaron coeficientes de validez más bien bajos o no generalizables, aunque los autores esperaban que también la estabilidad emocional fuera un predictor válido del desempeño. Respecto a los criterios organizacionales, solo el factor conciencia resultó ser un predictor valido y generalizable a través de todos los criterios, con coeficientes que varían de .20 (teórica) para los datos de personal a .23 para el éxito en el entrenamiento. Sin embargo, apertura a la experiencia y extraversión resultaron ser buenos predictores del éxito en la formación (.25 y .26 respectivamente, ambos coeficientes teóricos). En resumen, puede decirse que a pesar de los tamaños de validez moderados, estos resultados han sido muy positivos, sobre todo si tenemos en cuenta que los coeficientes de validez obtenidos serían superiores si se hubiese aplicado la mencionada corrección por la medición imperfecta del constructo (*composite score correction formula*), propuesta por Hunter y Schmidt (1990).

En el mismo año que el meta-análisis de Barrick y Mount (1991), se publicó otro estudio meta-analítico que utilizaba la clasificación de las escalas de personalidad según el modelo de los Cinco Grandes factores, el de Tett, Jackson y Rothstein (1991). El propósito de este estudio fue distinto que el de Barrick y Mount. Los autores pretendían demostrar que los resultados poco esperanzadores de las revisiones sobre personalidad, llevadas a cabo entre los años 50 y 80 del siglo XX, se deben principalmente a la falta de vinculaciones conceptuales entre el predictor y el criterio. En el caso del análisis de las propiedades predictivas de los Cinco Grandes, solo se utilizaron los estudios confirmatorios. Las muestras utilizadas fueron pequeñas y

varían de 2.302 personas el factor de extraversión a 280 personas para el factor de amigabilidad. Los resultados indicaron que todos los factores predicen desempeño laboral con los coeficientes de correlación que varían (en valores absolutos) desde .16 para extraversión hasta .33 para amigabilidad (ambos corregidos). Los coeficientes encontrados fueron superiores a los valores encontrados por Barrick y Mount. Además, la amigabilidad resultó ser mejor predictor del desempeño laboral que conciencia (.18). Los autores explican esta diferencia en la magnitud de los coeficientes encontrados aludiendo al hecho de que en el meta-análisis de Barrick y Mount se utilizaron los valores observados de los coeficientes encontrados, mientras que en su estudio se utilizaron los valores absolutos.

No obstante, cabe destacar que el meta-análisis de Tett y colaboradores (1991) presenta una serie de defectos. Primero, hay que hacer notar que los intervalos de confianza en todos los casos, con excepción del neuroticismo, incluían el cero, por lo que no pueden generalizarse. Este hecho puede deberse al tamaño pequeño de las muestras incluidas en el estudio de las capacidades predictivas de los *Cinco Grandes*. Comparando con el meta-análisis de Barrick y Mount (1991), el tamaño de las muestras utilizadas es de diez a veinte veces menor, dependiendo del factor. La mayor diferencia se da para los factores de amigabilidad y conciencia. Tett y colaboradores utilizaron una muestra total de solo 280 y 450 personas para amigabilidad y conciencia respectivamente, frente a 11.526 y 12.893 personas incluidas en la muestra de Barrick y Mount. Teniendo en cuenta que las fuentes bibliográficas de ambos meta-análisis fueron prácticamente iguales, esta diferencia en el tamaño muestral resulta sorprendente. Una posible explicación podría ser que Tett y colaboradores utilizaron para sus análisis solo los estudios confirmatorios, lo que pudo reducir substancialmente el tamaño de muestra. Sin embargo, incluso teniendo en cuenta esta diferencia, resulta extraño que en los análisis dedicados a la comparación de los estudios confirmatorios con los exploratorios se utilizó una muestra de 9,054 sujetos en total y un promedio de 197 sujetos en cada estudio incluido, mientras que el análisis de las propiedades predictivas de los *Cinco Grandes* y otros constructos de personalidad se hayan utilizado las muestras con un promedio de 100 personas en cada estudio y un número total de 8.156 sujetos. Tett y colaboradores no proporcionaron ninguna explicación para estas diferencias en los tamaños de las

muestras entre ambos análisis. Además, si tenemos en cuenta únicamente los cinco grandes factores, la muestra total es aún más pequeña, 5.236 sujetos. En segundo lugar, los autores tampoco explican los procedimientos utilizados para clasificar las escalas provenientes de distintos cuestionarios de personalidad en cinco grandes dimensiones. Por esa razón, es posible que las diferencias entre ambos meta-análisis de 1991 se deban a los diferentes procedimientos de clasificación de dichas escalas. En tercer lugar, es importante resaltar que Barrick y Mount resumieron las propiedades predictivas de los Cinco Grandes en cinco categorías ocupacionales, mientras que Tett y colaboradores presentan resultados a través de todas las ocupaciones. Por lo tanto, las diferencias entre ambos meta-análisis pudo ser el efecto moderador de la categoría ocupacional sobre la validez de las grandes dimensiones de personalidad. Todo ello pone de manifiesto que los resultados encontrados por Tett y colaboradores deben ser tratados con cautela, y que los resultados de ambos meta-análisis de 1991 no son comparables.

Sin embargo, a pesar de las diferencias entre los resultados obtenidos y los métodos utilizados por Hough y colaboradores (1990), Barrick y Mount (1991) y Tett y colaboradores (1991), estos tres meta-análisis se convirtieron en una base para la renovación del interés de los investigadores por las medidas de personalidad como predictores del desempeño. La importancia de estos estudios se debe al hecho de demostrar que los *Cinco Grandes* son una taxonomía muy útil a la hora de resumir los datos provenientes de múltiples estudios además de proporcionar, en el caso del estudio de Barrick y Mount, evidencia de que las dimensiones de personalidad se relacionan de manera distinta con diversos criterios organizacionales y a través de ocupaciones. Aunque las relaciones encontradas en este estudio pueden parecer modestas, también pueden considerarse útiles para la selección de personal, debido a que las medidas de personalidad muestran una validez añadida a la mayoría de los métodos de selección más comunes, como, por ejemplo, los tests de habilidades cognitivas (Rothstein y Goffin, 2006).

Las conclusiones generales provenientes de los tres meta-análisis arriba mencionadas son mucho más entusiastas que en el caso del meta-análisis de Schmitt y colaboradores (1984). Todas ellas dieron una base para una visión más optimista del potencial de la personalidad para predecir el desempeño laboral. Desde su publicación

se han llevado a cabo varios meta-análisis adicionales que relacionaron los *Cinco Grandes* factores con el desempeño laboral o alguna de sus facetas (p. ej., Anderson y Viswesvaran, 1998; C. M. Berry, Ones y Sackett, 2007; Hough, Ones y Viswesvaran, 1998; Hurtz y Donovan, 1998; Mount, Barrick y Stewart, 1998; Salgado, 1997, 1998a, 2002, 2003c).

Cabe destacar, pues, que la mayoría de los meta-análisis enumerados anteriormente se realizaron en los Estados Unidos y Canadá y las muestras de estudios que contienen se solapan en gran medida, con la excepción de dos meta-análisis de Salgado (1997, 1998a) que se hicieron con muestras únicamente europeas y no incluidas en los meta-análisis anteriores. Los resultados de estos meta-análisis demuestran que la validez de las medidas de personalidad puede generalizarse a países y zonas geográficas con características culturales y organizacionales diferentes. El primer meta-análisis (Salgado, 1997) sigue metodológicamente el modelo de Barrick y Mount (1991) y analiza 36 estudios con las muestras que iban desde 2.722 sujetos para apertura a la experiencia hasta 3.877 para la estabilidad emocional. Los coeficientes de validez encontrados muestran que estabilidad emocional y conciencia son los únicos factores con la validez predictiva generalizable a través de los tres tipos de criterios utilizados (valoraciones del desempeño, formación, y datos de personal, por ejemplo, rotación, absentismo, permanencia) y las cinco categorías ocupacionales: expertos, policías, puestos de gestión, puestos de venta y obreros cualificados, mostrando relaciones promedio de .19 y .25, respectivamente. También se encontró que la apertura a la experiencia y la amigabilidad eran predictores válidos del éxito en la formación, que la extraversión y apertura predecían el desempeño de los policías, además de que la extraversión era un predictor válido en el caso del desempeño de los obreros cualificados. Los resultados de esta investigación -realizada con estudios únicamente europeos- confirmaron las hipótesis originales de Barrick y Mount, que no había sido plenamente confirmadas en su estudio. Esencialmente, la estabilidad emocional resultó ser un predictor relevante del desempeño ocupacional, tal y como habían esperado. Un año más tarde, en un nuevo estudio meta-analítico, Salgado (1998a) confirmó los resultados obtenidos utilizando una muestra más amplia que incluía no solo las muestras civiles, sino también las muestras de ocupaciones militares. Conciencia y estabilidad emocional resultaron ser predictores válidos a

través de los estudios en muestras civiles y militares en conjunto con una validez de .23 promedio (ambos), sin embargo, estabilidad emocional fue mucho mejor predictor en el caso de las muestras de ocupaciones militares que de civiles (.45 frente a .19 respectivamente).

Unos años más tarde, Barrick, Mount y Judge (2001) realizaron un meta-análisis de varios estudios meta-analíticos anteriores. Sus resultados reflejan la capacidad predictiva de personalidad demostrada a través de distintos países e incluso continentes, al igual que ocupaciones y criterios. Este meta-análisis confirmó que los *Cinco Grandes* se relacionan de manera diferente con distintos criterios laborales y a través de ocupaciones. De esta forma, dependiendo del objetivo de estudio, cada uno de los factores puede ser un predictor muy útil. La conciencia es un predictor generalizable del desempeño a través de todos los criterios (.27), pues predice tanto el trabajo en equipo (.27) como el éxito en la formación (.27), las valoraciones de los supervisores (.31) y los criterios objetivos del desempeño (.23). En cuanto a la estabilidad emocional, esta predice desempeño general (.13), pero también los criterios más específicos como el trabajo en equipo (.22). Extraversión resultó ser un buen predictor del trabajo en equipo (.16), del éxito en la formación (.28), del desempeño de policías (.12) y de gerentes (.21). Asimismo, la amigabilidad mostró una relación significativa y generalizable con el trabajo en equipo (.34). Por último, la apertura resultó a su vez predecir el éxito en la formación (.33).

Un año más tarde, Salgado (2002) publicó un meta-análisis nuevo donde se investigó la relación entre los cinco grandes factores de personalidad y las conductas contraproductivas en el trabajo. Al igual que en meta-análisis anteriores, las escalas de diferentes inventarios de personalidad fueron agrupadas según el modelo del los *Cinco Grandes* pero esta vez, en lugar del desempeño laboral como criterio se utilizaron conductas tales como absentismo (ausencias y pérdidas de tiempo), los accidentes, las conductas desviadas (robos, admisiones de robos, problemas disciplinarios, etc.), y la rotación en el empleo (tasa voluntaria de despido y abandono del trabajo). Para las conductas desviadas, las relaciones generalizables resultaron ser con la amigabilidad (-.20) y con la conciencia (-.26). A su vez, todas las dimensiones de personalidad resultaron predecir la falta de la rotación en el empleo, con los coeficientes que varían de .11 para la apertura hasta .25 para la estabilidad emocional. Ni los accidentes, ni el

absentismo fueron predichos en este estudio por las grandes dimensiones de personalidad.

Posteriormente, en lo que se refiere a las conductas contraproductivas, los resultados de Salgado (2002) fueron parcialmente confirmados por el meta-análisis de C. M. Berry, Ones y Sackett (2007). En este estudio, se investigó la relación de los Cinco Grandes con dos tipos de conductas contraproductivas: organizacionales e interpersonales. Las primeras implican comportamientos negativos para las organizaciones tales como robo, destrucción de maquinaria, abuso de alcohol y drogas, un ritmo lento de trabajo, hacer pública información confidencial, etc. Las segundas abarcan las conductas dirigidas a las personas tales como violencia física, agresiones verbales, rumorear y robar a los compañeros. En este meta-análisis al igual que en el estudio de Salgado, la amigabilidad y la conciencia resultaron ser predictores válidos y generalizables de conductas desviadas organizacionales, además de estabilidad emocional, alcanzando los coeficientes de -.32, -.42 y -.23, respectivamente. Al mismo tiempo, las conductas desviadas interpersonales fueron predichas por los mismos factores, en cambio, los coeficientes de validez fueron algo inferiores: -.36, -.19 y -.20, respectivamente.

En términos generales, tal y como destaca Salgado (2003c), hasta el meta-análisis de Barrick y colaboradores (2001), la acumulación de estudios sobre personalidad se hacía principalmente clasificando las escalas de personalidad provenientes de diferentes inventarios, según el modelo de los *Cinco Grandes*. Sin embargo, esta clasificación puede presentar una serie de problemas. Entre otros, los investigadores pudieron cometer errores durante la clasificación, incluso a pesar de un elevado grado de acuerdo entre los jueces. Por ejemplo, en el meta-análisis de Hough y colaboradores (1990), la correlación promedio entre las escalas, que supuestamente deben medir el mismo constructo, varía de .33 a .46. El error de clasificación es uno de los posibles errores artifactuales (Hunter y Schmidt, 1990) que pueden afectar a la estimación de la validez verdadera y a la varianza de los *Cinco Grandes*. Otro error artifactual consiste en que las escalas que miden solo una parte del constructo son agregadas y se calcula la media de los coeficientes de correlación de las mismas con el criterio de interés, tratándolas como las medidas con la validez de constructo completa. Sin embargo, esta es una medición del constructo imperfecta y este error

artifactual también puede subestimar la validez de los *Cinco Grandes*. Hay dos caminos posibles para evitar el error de medición imperfecta del constructo. Por un lado, se puede aplicar la corrección de la validez por la medición imperfecta del criterio (*composite score correction formula*) propuesta por Hunter y Schmidt. Aun así, esta corrección se ha utilizado solo en los meta-análisis de Mount y Barrick (1995) y Salgado (1997, 1998a). Por otro lado, este tipo de error puede evitarse utilizando únicamente las medidas desarrolladas bajo el modelo de los *Cinco Grandes*. Es necesario resaltar que hasta el año 1991 existían solo tres inventarios de personalidad basados en el modelo de los Cinco Grandes. En 2003, Salgado aseguró que, debido a que en aquel entonces existía ya más de una quincena de cuestionarios diferentes desarrollados en el marco de este modelo, cabía la posibilidad de resumir de forma cuantitativa los resultados de los estudios basados únicamente en los inventarios de personalidad creados para medir los *Cinco Grandes*. Por lo tanto, comparó la validez de los cuestionarios basados en los *Cinco Grandes* con la de los inventarios de personalidad no basados en éstos (pero clasificados según este modelo) para predecir el criterio de las valoraciones del desempeño. Su estudio abarcó 139 muestras independientes de las que 74 fueron muestras norteamericanas con coeficientes de validez calculados para las escalas de los inventarios de personalidad no desarrollados en el marco de los *Cinco Grandes*, y 17 fueron muestras europeas del mismo tipo. De las muestras restantes, 26 fueron norteamericanas en las que se aplicaron medidas específicas de los *Cinco Grandes* y otras 22 fueron europeas del mismo tipo. Los resultados de este meta-análisis indicaron que el uso de las medidas desarrolladas específicamente para medir las cinco grandes dimensiones de personalidad mejora substancialmente los coeficientes de la validez predictiva. La validez operacional encontrada para el factor de conciencia -utilizando los inventarios basados en el modelo de los *Cinco Grandes*- fue de .28, mientras que en el caso de los inventarios no basados en los *Cinco Grandes* fue de .18, y además el intervalo de confianza en el segundo caso incluía cero, por lo que este resultado no es generalizable. En el caso de la estabilidad emocional, la validez operacional de las medidas basadas en el modelo de los *Cinco Grandes* fue de .16, mientras que para los inventarios no basados en los *Cinco Grandes* fue de .05, pero otra vez el intervalo de confianza incluía cero, por lo que este coeficiente tampoco se puede generalizar. Sorprendente es el hecho que en el caso de medidas no

basadas en los *Cinco Grandes*, la validez de conciencia y estabilidad emocional no pudieran generalizarse. Este resultado es inconsistente con los meta-análisis anteriores (Barrick y Mount, 1991; Barrick et al., 2001; Salgado, 1997), en los que también se utilizó la clasificación de las escalas de los inventarios de personalidad no creados específicamente para medir los *Cinco Grandes*. Sin embargo, la validez predictiva de estos dos factores resultó ser generalizable. El autor explicó este hallazgo alegando que, aunque la validez de la conciencia y la estabilidad emocional para predecir las valoraciones del desempeño sea generalizable, en el caso de los inventarios no basados en el modelo de los *Cinco Grandes* existe una varianza artifactual añadida a los coeficientes observados. Esta varianza artifactual puede producir, en algunos casos, el efecto de la validez operacional negativa, en concreto, en el caso de las ocupaciones en las que la validez de la conciencia y la estabilidad emocional sea menor, o en el caso de muestras pequeñas. El análisis de la validez predictiva de ambos tipos de medidas (basadas y no basadas en los Cinco Grandes) resultó, en conjunto, en coeficientes de validez de .26, .13, .08, .10 y .11 para los factores de conciencia, estabilidad emocional, apertura a la experiencia, amigabilidad y extraversión, respectivamente. Por lo tanto, en el caso de la conciencia y la estabilidad emocional los resultados son casi idénticos a los del meta-análisis de los meta-análisis realizado por Barrick y colaboradores (2001) (.23 y .13 frente a .27 y .13). A su vez, en el caso de los factores restantes, los resultados son también muy similares.

En resumen, los meta-análisis de personalidad existentes confirmaron que la personalidad predice no solo el desempeño general, sino también el éxito en la formación, los criterios objetivos del desempeño y las conductas contraproductivas. Sin embargo, a lo largo de los años, muchos otros estudios meta-analíticos han comprobado que los *Cinco Grandes* tienen también una relación positiva y generalizable con muchos más criterios relacionados con el trabajo, como puede ser *progresar* (*getting ahead*) (J. Hogan y Holland, 2003), *llevarse bien* (*getting along*) (J. Hogan y Holland, 2003), el desempeño de tarea (Dudley, Orvis, Lebiecki y Cortina, 2006; Hurtz y Donovan, 2000) o con distintas dimensiones del desempeño contextual, entre otros, altruismo, dedicación al trabajo, facilitación interpersonal y obediencia generalizada (*generalizad compliance*) (Borman, Penner, Allen y Motowidlo, 2001; Dudley et al., 2006; Hurtz y Donovan, 2000; LePine, Erez y Johnson, 2002). A su vez, se

demostró que los *Cinco Grandes* predicen el aprendizaje y la adquisición de habilidades (Colquitt, LePine y Noe, 2000) y que son un indicador de eficacia gerencial y de promoción de gerentes (Hough et al., 1998). El meta-análisis de Mol, Born, Willemsen y Van Der Molen (2005) confirmó que la personalidad predice el desempeño de los expatriados. Zhao y Seibert (2006) hallaron que la personalidad es un predictor válido de la capacidad emprendedora. Otros estudios se han centrado en demostrar la relación entre los *Cinco Grandes* y el establecimiento de metas y la autoeficacia (Judge y Ilies, 2002) o, por el contrario, el aplazamiento (Steel, 2007). Las variables de personalidad resultaron también ser buenos predictores de creatividad e innovación (Feist, 1998; Hough, 1992). Además, los *Cinco Grandes* han demostrado su utilidad a la hora de predecir varios criterios de liderazgo: aparición de liderazgo, eficacia de liderazgo (Judge, Bono, Ilies y Gerhardt, 2002) y liderazgo transformacional (Bono y Judge, 2004). Peeters, Van Tuijl, Rutte y Reymen (2006) mostraron a su vez que la personalidad es un predictor válido del desempeño del grupo. También se halló una relación positiva y generalizable entre los *Cinco Grandes* y las actitudes hacia el trabajo tales como la satisfacción laboral (Judge, Heller y Mount, 2002) y la satisfacción con la carrera profesional (Ng, Eby, Sorensen y Feldman, 2005). Finalmente, los *Cinco Grandes* mostraron también capacidades predictivas en cuanto a los accidentes laborales (Clarke y Robertson, 2005).

Estos resultados confirmaron definitivamente que el modelo de los *Cinco Grandes* es una taxonomía sumamente útil para acumular y comparar los datos de diferentes estudios e incluso diferentes países y que predice además un amplio abanico de criterios organizacionales. Finalmente, después de un largo periodo de escepticismo de los psicólogos organizacionales, la personalidad ha recuperado su interés y ha vuelto a ser reconocida como un concepto relevante para las organizaciones (Rothstein y Goffin, 2006). Hoy en día, los inventarios de personalidad se utilizan en los procesos de selección debido a sus propiedades psicométricas demostradas que, a pesar de ser moderadas, aportan una validez añadida a otros métodos comúnmente utilizados en la selección de personal (como, por ejemplo, los tests de habilidades cognitivas) (Barrick y Mount, 2005; Salgado, 1998a). Al mismo tiempo, es importante destacar también que el impacto adverso de las medidas de personalidad es mínimo o casi nulo en comparación con la mayoría de los métodos de

selección (Hough et al., 2001; Ones y Anderson, 2002). Este es otro punto importante a favor del uso de los tests de personalidad en los procedimientos de selección y de promoción en las organizaciones.

CAPÍTULO 2

EL DEBATE SOBRE LA AMPLITUD Y LA FIDELIDAD Y LA PREDICCIÓN DEL DESEMPEÑO OCUPACIONAL

La popularidad del modelo de los cinco grandes factores en la evaluación de la personalidad ha originado un intenso debate que abarca entre otras la cuestión de la exhaustividad de los factores. Algunos investigadores (p. ej., Ashton, 1998; Hough, 1992; Kanfer, Ackerman, Murtha y Goff, 1995; Paunonen, 1998; Paunonen, Jackson, Trzebinski y Forsterling, 1992; Paunonen, Rothstein y Jackson, 1999; Tett, Jackson, Rothstein y Reddon, 1994) han cuestionado si los cinco grandes factores de personalidad (*broad measures*) son suficientemente robustos como para proporcionar una predicción superior a la que otorgan las numerosas variables de personalidad (*narrow measures*) denominadas rasgos del orden inferior, subdimensiones, facetas, clusters, o conjuntos de ítems homogéneos (*Homogeneus Item Clusters – HICs*), más específicos que los *Cinco Grandes.* Esta controversia se conoce en la literatura científica bajo la denominación: "dilema amplitud-fidelidad" (*bandwidth-fidelity dilemma*).

El origen del concepto de dilema amplitud-fidelidad

El dilema amplitud-fidelidad es un término originalmente propuesto por Cronbach y Gleser (1957) con el fin de describir las propiedades psicométricas de los tests cognitivos, aunque posteriormente también se aplicó a los inventarios de personalidad. Cronbach y Gleser postularon que en la medición de los constructos psicológicos siempre se produce una incompatibilidad (*trade-off*) entre la amplitud de la medida y la fidelidad de la misma, de modo que: (a) los constructos amplios predecirían los criterios amplios con una validez moderada y (b) los constructos más estrechos predecirían los criterios específicos con una mayor validez. Aplicando estos supuestos a los inventarios de personalidad nos encontraríamos con la situación en la que los factores de personalidad amplios, como los Cinco Grandes, serían útiles para predecir criterios amplios, aunque su validez predictiva sea menor que en el caso de las medidas más específicas, las facetas, que predecirían criterios más delimitados, pero con una mayor exactitud (John y Robins, 1993; McAdams, 1992).

Los comienzos del debate

Cattell y colaboradores (1970), comparando su modelo de 16 factores con los modelos alternativos que abarcan un número de factores más reducido (p. ej., el de Eysenck o el de Norman) concluyeron que éstos operan en un nivel de abstracción superior, por lo que pierden una parte importante de información. También otros autores se han reafirmado en la convicción de que la mayor precisión de medida necesita rasgos más específicos, a pesar de que las dimensiones amplias posean la ventaja de proporcionar una estructura de personalidad fácilmente interpretable (Briggs, 1989, 1992; Buss, 1989; Mershon y Gorsuch, 1988).

Aunque intuitivamente estas premisas pueden parecer ciertas, algunos investigadores han cuestionado su veracidad en el caso de las medidas de personalidad, dando origen al debate mencionado anteriormente sobre si son mejores predictores del desempeño laboral los factores amplios o bien sus facetas. Ones y Viswesvaran (1996) argumentaron que las medidas amplias normalmente tienen una

consistencia interna superior a la de las medidas estrechas, por lo que su validez de criterio también es superior. Además, según estos autores, la mayoría de los criterios del desempeño laboral son factorialmente complejos, por lo que se requieren predictores igual de complejos para conseguir una validez máxima. Al mismo tiempo, opinaron que el compuesto de integridad, que correlaciona altamente con tres de los cinco grandes factores (conciencia, estabilidad emocional y amigabilidad), es todavía mejor predictor del desempeño laboral que los factores subyacentes. Según Ones y Viswesvaran, esto demostraba que las escalas más amplias son más válidas y podía ser un argumento a favor de la existencia de un factor general de personalidad comparable con el factor *g* de las medidas de habilidad mental general.

J. Hogan y Roberts (1996) argumentaron a su vez que la cuestión principal en este debate es el emparejamiento de la amplitud del predictor y del criterio. Schneider, Hough y Dunette (1996) también plantearon que los predictores deberían ser equiparables con los criterios en los términos de la especificidad, y que los mejores coeficientes de la validez de criterio se obtienen cuando los investigadores usan el enfoque orientado al constructo para emparejar el predictor y el criterio.

La cuestión de la amplitud del criterio

Así pues, J. Hogan y Roberts (1996) y Schneider y colaboradores (1996) han añadido al debate sobre el intercambio entre amplitud y fidelidad, la cuestión de la amplitud del criterio del desempeño laboral. También en este caso hay diversidad de opiniones sobre si es mejor predecir un criterio global de desempeño laboral (*overall job performance*), o si es preferible predecir criterios de desempeño más específicos, por ejemplo, el desempeño de tarea o el contextual. Lo que dificulta un consenso sobre este aspecto es la falta de acuerdo entre los investigadores sobre la taxonomía de las medidas de desempeño. Sin embargo, a lo largo de los años se han desarrollado diferentes propuestas que pueden ser adaptadas y utilizadas (p. ej., Borman y Motowidlo, 1993; Campbell, 1990; Campbell, McCloy, Oppler y Sager, 1992; Viswesvaran, Schmidt y Ones, 2005). Además, aunque no existiese una única taxonomía de las dimensiones del desempeño laboral, hoy en día tanto los investigadores como los profesionales están de acuerdo en que el desempeño laboral

es una variable multidimensional y existe evidencia teórica y empírica de que distintas variables de personalidad pueden ser más o menos relevantes en la predicción de distintos criterios del desempeño (p. ej., Barrick et al., 2001; C. M. Berry et al., 2007; Dudley et al., 2006; Salgado, 2002).

Los tres principales grupos de posiciones en el debate

Optar por medidas de personalidad amplias (*broad*) frente a estrechas (*narrow*) y optar por medidas del desempeño globales (*broad*) frente específicas (*narrow*), da lugar a tres grupos de posiciones principales en el debate sobre la incompatibilidad (*trade-off*) entre la amplitud y la fidelidad en los inventarios de personalidad: (1) aquellos autores que sostienen que las medidas amplias de personalidad predicen mejor los criterios amplios que las medidas estrechas de personalidad y, además, predicen igualmente bien los criterios específicos (*narrow criteria*) (Chapman, 2007; Mount y Barrick, 1995; Ones y Viswesvaran, 1996); (2) una segunda posición, opuesta a la anterior, sostiene que las medidas estrechas de personalidad son mejores predictores de los criterios específicos, y añaden varianza a las medidas globales de personalidad para explicar las medidas globales del criterio (Ashton, 1998; Christiansen, 2008; Paunonen et al., 1999; Tett y Christiansen, 2007; Tett, Steele y Beauregard, 2003); por último, (3) una tercera posición respaldada por aquellos autores que sostienen que debe haber una alineación entre predictores y criterios, de tal modo que los criterios amplios (globales) serán mejor predichos por las dimensiones globales y los criterios específicos por las subdimensiones o las facetas de personalidad (*narrow measures*) (J. Hogan y Roberts, 1996; Moberg, 1998; Schneider et al., 1996). Hasta el momento, no existe ningún estudio que haya examinado completamente estas cuestiones, no obstante, existen algunos estudios primarios y meta-analíticos recientes que las han examinado parcialmente.

El dilema amplitud-fidelidad reflejado en los estudios primarios

Uno de los primeros investigadores en examinar esta cuestión ha sido Asthon (1998). Este investigador cuestionó la existencia de un amplio superfactor de integridad en las medidas de personalidad, argumentando que si las escalas de conciencia, amigabilidad y estabilidad emocional y la integridad estuviesen realmente relacionadas, los inventarios de personalidad deberían contener un amplio factor de integridad, sin embargo, esto habitualmente no ocurre. Además, aseguró que si hubiese algún factor general, la validez de los tests de integridad se debería a la varianza compartida con los Cinco Grandes, sin embargo las investigaciones demuestran que no es así. En su estudio encontró que las que él denomina *medidas amplias de personalidad* están ligeramente menos correlacionadas con las conductas contraproductivas que las que él denomina "medidas estrechas". En concreto, encontró que la contraproductividad se relacionaba de manera más significativa con dos de los cinco grandes factores, amigabilidad (-.21) y conciencia (-.22), y con dos escalas: tendencia a tomar riesgos (.30) y responsabilidad (-.40). Mediante el método de regresión múltiple ponderada creó un compuesto con los dos factores alcanzando un coeficiente de validez de -.38 (no ajustado), mientras que el compuesto de ambas facetas resultó en un coeficiente de -.45 (no ajustado). Sin embargo, como surgiere el mismo autor debido a la alta correlación entre ambos compuestos (.76), la diferencia entre las correlaciones a favor de las medidas específicas no es estadísticamente significativa. Esta alta correlación entre los compuestos superaría .90 si fuese corregida por la fiabilidad, lo que nos indica que los constructos analizados son equiparables. Por lo tanto, realmente estamos hablando de los mismos constructos. Cabe destacar que las escalas mencionadas, junto con otras escalas utilizadas en este estudio, cargan altamente en más de un factor por lo que la relación entre las mismas también es alta. De este modo, la tendencia a tomar riesgos tiene una relación de -.44 con conciencia y de .41 con apertura a la experiencia, mientras que la responsabilidad se relaciona con amigabilidad (.56) y con conciencia (.44). Esto nos indica que nos enfrentamos a variables compuestas más que a facetas. Las facetas, aunque pueden tener cargas secundarias, éstas deben ser mínimas, mientras que en este caso ambas son

relativamente altas. Además, como otra debilidad del estudio hay que resaltar que se utilizaron medidas diferentes para medir los factores y las facetas (en el caso de las últimas, debido a que muchas de las escalas utilizadas en este estudio cargan significativamente en más de un factor, deberían considerarse variables compuestas). Mientras que los factores provenían del listado de adjetivos marcadores (*adjective markers*) de Goldberg (1992), las escalas formaban parte del *Jackson Personality Inventory* (JPI; D. N. Jackson, 1976, 1994) (15 escalas) y del *Personality Research Form* (PRF; D. N. Jackson, 1984) (1 escala).

El hecho de utilizar variables compuestas como facetas de los Cinco Grandes, al igual que el uso de diferentes instrumentos para la medición de los factores y de las facetas, como en el caso del estudio de Ashton (1998), conduce a una confusión relevante, porque en realidad no sabemos a qué amplitud de constructo nos estamos enfrentando. El hecho de llamar a un conjunto de ítems "una escala" no lo convierte en una faceta, al igual que llamar a un conjunto de escalas "un factor" no implica la existencia de una variable latente. Para constatar esto con toda claridad necesitamos disponer de la matriz de datos y conocer su estructura factorial. No se puede olvidar que son las facetas las que *dependen del factor* -y no al revés- y una parte de la varianza del criterio explicada por cada faceta se debe al hecho de que ésta esta afectada por una o más variables latentes (los factores). De acuerdo con la teoría clásica de la medida y del análisis factorial (Crocker y Algina, 1986) esta parte de varianza que la faceta debe a los factores se denomina *comunalidad* o *varianza común.* Aparte de la *varianza común* que la faceta comparte con otras facetas pertenecientes al mismo factor, la validez de la faceta depende también de su varianza única. Esta parte de varianza se denomina *unicidad* e indica hasta qué punto las variables latentes que influyen en la faceta son insuficientes para explicar toda su varianza. Teóricamente, la unicidad puede dividirse en dos partes: la *varianza específica* y la varianza *de error*. La *varianza específica* sería esta proporción de la puntuación verdadera (*true score*) de la faceta que no está relacionado con la puntuación verdadera de ninguna otra faceta. La *varianza de error* sería el componente de error que tiene lugar en cualquier medida. Por lo tanto, el verdadero poder predictivo de la faceta es solo el que corresponde a su varianza específica. Utilizando medidas diferentes para medir factores y facetas es imposible estimar la varianza debida a la

comunalidad, al error y calcular el valor exacto de la *varianza específica*. El mismo problema ocurre cuando se utilizan los compuestos de facetas para conseguir una mejor predicción. En este caso, tampoco se tiene en cuenta que el factor no es una simple suma de facetas, sino que es una variable latente que transfiere su varianza a las facetas causando que éstas se relacionen entre sí.

Para concluir, a la hora de estudiar el verdadero poder predictivo de las facetas, la varianza debida al factor y al error debería ser descontada de la varianza total de la faceta, algo que no se ha hecho en el estudio de Ashton.

Finalmente, cabe resaltar que el método de regresión múltiple, tal y como se ha utilizado en el estudio de Ashton (1998), no es adecuado para resolver el dilema amplitud-fidelidad. Esto se debe a que la regresión múltiple en sí no permite descomponer la varianza de los elementos del modelo en comunalidad, *varianza específica y varianza del error*, por lo que no sabemos qué parte de esta varianza es la responsable de la relación encontrada. Además, esta observación es cierta tanto en los casos cuando se encuentra la validez añadida de las facetas, como de los factores. Para poder utilizar de forma correcta la regresión múltiple, hay que utilizar los valores puros en cada caso, es decir, los coeficientes de validez predictiva que se deben únicamente a la *varianza del factor*, o a la *varianza específica* de las facetas. En cambio, como veremos adelante, en ninguno de los estudios existentes se ha tenido en cuenta la dependencia de las facetas del factor. La mayoría de los investigadores tratan a la varianza de una faceta como equivalente a su *varianza específica*, sin tener en cuenta otros elementos de la varianza, comunalidad y error. Por lo tanto, el método de regresión múltiple de forma no correcta se aplica a los coeficientes sesgados. Además, la elección de las facetas en la mayoría de los casos es arbitraria, lo que impide las comparaciones de los resultados entre diferentes estudios. Esta arbitrariedad tiene su origen en la falta de una taxonomía universal de facetas. La existencia de dicha taxonomía ayudaría a los científicos a acumular datos sobre el poder predictivo de personalidad en el menor nivel de abstracción. Éste es un inconveniente al que se enfrentan muchos investigadores que estudian las cuestiones relacionadas con el dilema amplitud-fidelidad, por lo que las facetas medidas en estos estudios son diferentes prácticamente en cada caso.

Otro estudio importante para el tema que nos ocupa fue realizado por Paunonen (1998), quien siguiendo la línea de Ashton (1998) comparó en dos estudios las capacidades predictivas de los factores y rasgos específicos en cuanto a un conjunto de criterios que denominó “de importancia social” (pág. 540). Estos criterios incluían: los hábitos de fumar, la nota media académica y las evaluaciones de atracción física de los compañeros, y fueron elegidos para representar variables multi-facéticas vinculadas posiblemente con más de una dimensión de personalidad. En ambos estudios encontró que las facetas tienen capacidades predictivas superiores a los factores, lo que se explica con la pérdida de *varianza específica* relevante para los criterios en cuestión. Esta conclusión se basaba en el hecho de que aunque los grandes factores mostraron validez añadida sobre las escalas específicas y viceversa, la validez añadida de las facetas fue mayor que la de los factores. Sin embargo, cabe destacar que en ambos estudios se utilizaron medidas distintas para los rasgos específicos y para los factores. Mientras que las escalas consideradas por el autor *facetas* provenían del PRF (D. N. Jackson, 1984) (el primer estudio) y del JPI (D. N. Jackson, 1976, 1994) (el segundo estudio), los factores han sido determinados mediante el uso del *NEO-Five Factor Inventory* (NEO-FFI; Costa y McCrae, 1992b). Vemos por lo tanto que Paunonen, al igual que Ashton (1998), considera que las escalas de PRF y de JPI son las facetas puras, mientras que estudiando la estructura factorial de ambos inventarios se observa que en realidad son variables compuestas, al tener altas cargas factoriales en más de un factor. Además, al utilizar cuestionarios diferentes no podemos comparar la amplitud de los constructos evaluados. Sin embargo, incluso sin tener en cuenta estas limitaciones, se utilizaron medidas más largas para las escalas específicas que para los factores (las escalas del PRF son de 16 ítems, las del JPI de 20 ítems, mientras que el NEO-FFI mide cada factor con solo 12 ítems), lo que hace que sean también más fiables. Acorde con la sugerencia de los autores del dilema amplitud-fidelidad, Cronbach y Gleser (1957), si una medida de un único constructo tuviese un número de ítems igual o superior que otra medida multidimensional, la primera tendría una validez superior a la segunda. A su vez, tal y como sugirió Chapman (2007), en la medición de los Cinco Grandes esto normalmente no tiene lugar debido a que el número de ítems que mide cada faceta suele ser muy inferior al número total de ítems que mide el factor. Los errores metodológicos antes mencionados podrían ser la razón

de las propiedades predictivas superiores de las escalas específicas halladas por Paunonen (1998). Además, en este estudio también se utilizó el método de análisis de regresión múltiple de forma incorrecta y no adecuada para resolver las cuestiones aquí planteadas, ya que previamente no se llevó a cabo la residualización de la varianza.

Moberg (1998) al igual que Aston (1998) y Paunonen (1998) se mostró partidario de las medidas estrechas, partiendo del supuesto de que utilizando las medidas amplias de personalidad se omite la información específica que contienen las facetas, oscureciendo de este modo las relaciones conceptuales entre el predictor y el criterio. Para medir los factores y las facetas utilizó el cuestionario NEO PI-R (Costa y McCrae, 1992b), a su vez, para estimar si las facetas o los factores son mejores predictores de la estrategia del conflicto se basó en el método de regresión múltiple. Encontró que la combinación de las facetas vinculadas conceptualmente con la estrategia del conflicto (*conflict strategy*) es mejor predictor de este criterio que un compuesto de varios factores. Analizando el poder predictivo de los factores y las facetas correspondientes a cada factor por separado, Moberg concluyó que en ninguna de las correlaciones analizadas el poder predictivo de los factores era superior al de las facetas. En todos los 20 casos (cuatro diferentes estrategias de conflictos multiplicadas por cinco factores) se encontró apoyo para los supuestos iniciales: en dos casos los compuestos de facetas explicaron el mismo porcentaje de varianza que los factores, en tres casos explicaron un poco más de la varianza que los factores, en siete casos explicaron dos veces más de la varianza que los factores, y en ocho casos tuvieron una relación significativa con el criterio cuando los factores no mostraron ninguna. Estos resultados Moberg interpretó a favor de la necesidad de emparejamiento entre el predictor y el criterio, debido al aumento de las propiedades predictivas del predictor. Sin embargo, cabe destacar que el autor no presenta los intervalos de confianza, por lo que no se puede estimar el grado de solapamiento que hay entre los mismos, algo que influye en la generalizabilidad de los resultados. Además, utilizando el método de regresión múltiple, no se estimó previamente el verdadero valor predictivo de las facetas (las facetas no se trataron como un constructo puro). En consecuencia, la varianza total de la faceta ha sido tratada como su varianza específica, pese a ser erróneo este supuesto.

Los estudios de Ashton (1998), Paunonen (1998) y Moberg (1998) han aumentado el ya existente interés por la relación entre la amplitud de las medidas de personalidad y sus capacidades predictivas. Sin embargo, ninguno de estos trabajos se centró en el criterio más relevante para los psicólogos organizacionales: el desempeño laboral. Entre los estudios primarios que tratan el tema del dilema amplitud-fidelidad utilizando el desempeño laboral como criterio, cabe destacar el trabajo de Salgado, Rumbo, Santamaría y Rodríguez-Losada (1995). Este grupo de investigadores comparó el poder predictivo de los factores primarios del cuestionario 16 PF (Cattell et al., 1970) tratados de forma independiente, y agrupados según el modelo de los Cinco Grandes. Las puntuaciones alcanzadas en cada uno de los factores primarios y en cada uno de los Cinco Grandes han sido relacionadas con el criterio del desempeño en el sector de banca. Para conocer la capacidad predictiva global de cada una de las estructuras estudiadas -5 grandes factores frente a los 16 factores primarios- se llevó a cabo el análisis de regresión múltiple. Aunque el tamaño del coeficiente de correlación múltiple fue similar en ambos casos, solo la regresión derivada del modelo de los Cinco Grandes resultó ser significativa. Salgado y colaboradores (1995) encontraron que neuroticismo resultó ser buen predictor del desempeño (-.22), sin embargo, ninguno de los factores primarios que formaban su parte lo fueron. Con todo ello, los autores del estudio concluyen que los resultados encontrados no apoyan la hipótesis de que las facetas podrían tener un poder predictivo superior a los factores del orden superior, por lo menos, cuando el criterio a predecir es el rendimiento en el trabajo. Sin embargo, este estudio también presenta una serie de limitaciones. En primer lugar, cabe destacar el pequeño tamaño de la muestra (N=93) utilizada, que condiciona la generalización del hallazgo. En segundo lugar, el cuestionario analizado (el 16 PF) no ha sido creado para reflejar el modelo de los Cinco Grandes, y el proceso de agrupar las escalas de un cuestionario creado dentro de un enfoque diferente al modelo de los cinco grandes factores según el cual puede dar lugar a un error de clasificación que sería uno de los posibles errores artifactuales (Hunter y Schmidt, 1990) que pudiesen afectar a la estimación de la validez teórica y de la varianza de los factores. Este error puede ocurrir incluso a pesar de seguir un estricto procedimiento, y a pesar de un elevado grado de acuerdo entre los jueces (R. T. Hogan, Hogan y Roberts, 1996). La única manera de evitarlo sería utilizar los instrumentos que a priori han sido

desarrollados como medidas de este modelo. Finalmente, cabe destacar que en el estudio de Salgado y colaboradores, los cinco factores fueron tratados conceptualmente como los compuestos de varias facetas (en este caso de los factores primarios del 16 PF). Ahora bien, el factor es más que un simple compuesto, es una variable subyacente que influye en las facetas, algo que en este caso no se puede constatar sin conocer las cargas factoriales de cada uno de los factores primarios del 16 PF en los Cinco Grandes correspondientes.

El siguiente estudio primario que examinó temas relacionados con el dilema amplitud-fidelidad utilizando como el criterio el rendimiento laboral es el de Stewart (1999). En él se analizó la validez predictiva del factor conciencia y de dos componentes del mismo (logro y orden) para predecir el criterio de desempeño laboral dividido en dos etapas específicas. La primera etapa fue denominada *de transición* y ocurría cuando los empleados empezaban el nuevo trabajo, o cuando sus obligaciones y responsabilidades cambiaban de manera significativa. La segunda etapa, llamada *de mantenimiento*, se da cuando los empleados ya tienen el conocimiento, las capacidades, y las habilidades asociadas con el puesto que ocupan. El factor de conciencia fue medido mediante el *NEO Personality Inventory Revised* (NEO PI-R; Costa y McCrae, 1992b) (48 ítems). Las facetas formaban parte de la medición de este factor junto con otras cuatro facetas que para los análisis detallados no se las tuvo en cuenta, y cada una se medía con ocho ítems. Se encontró una relación entre el desempeño general y el factor de conciencia de .16. En cuanto a las facetas, la del orden resultó ser un buen predictor del desempeño general (.16), comparable con el factor de conciencia. Sin embargo, la relación del desempeño con el logro no fue significativa. En ambos grupos de trabajadores (nuevos y veteranos) se encontró una relación positiva y consistente entre el desempeño y el factor de conciencia. Sin embargo, la relación entre la faceta de orden y el desempeño fue más alta en la muestra de trabajadores en la etapa de transición (.27) que en la muestra de los trabajadores en la etapa de mantenimiento (.06). A su vez, el logro fue mejor predictor del desempeño en la etapa de mantenimiento (.22) que en la etapa de transición (-.01). Por consiguiente, un modelo de regresión que incluía tanto el factor de conciencia como sus componentes fue comparado con un modelo basado solamente en el factor de conciencia. Se encontró un cambio significativo en el R^2 entre dos modelos, lo que se interpretó como

un índice de la existencia de validez añadida. La faceta de orden ha explicado un 4% de la varianza en el desempeño de los trabajadores en la etapa de transición por encima del factor de conciencia. El logro mostró una validez añadida en la muestra de trabajadores en la etapa de mantenimiento de un 3%. Por lo tanto, Steward concluyó que aunque el factor de conciencia tenía una relación consistente con el desempeño laboral a través de ambas etapas del desempeño, las facetas de orden y de logro son mejores predictores del desempeño cuando se analizan estas etapas por separado, ya que muestran relaciones diferentes en cada una de ellas. Sin embargo, debido a que en este estudio se trató a las facetas como variables independientes y sin estimar su varianza única habría que volver a comparar el poder predictivo de conciencia frente al de las facetas utilizando solo su varianza específica para poder sacar las conclusiones definitivas.

Tett y colaboradores (2003), al igual que algunos investigadores que realizaron sus estudios con anterioridad (p. ej., Ashton, 1998; Moberg, 1998; Paunonen, 1998), supusieron que el hecho de dirigir la atención hacia un contenido más específico aumenta la utilidad de las medidas de personalidad, a pesar de la reducción en la amplitud. En su opinión, el uso de las medidas amplias oscurece las relaciones significativas entre las medidas más específicas y el criterio. Por lo tanto, el hecho de basarse únicamente en los factores puede impedir una diagnosis correcta sobre los vínculos reales entre personalidad y conducta laboral. Para verificar su teoría llevaron a cabo dos estudios primarios cuyo objetivo era comparar las propiedades predictivas de los factores amplios en comparación con las de sus facetas respecto a dos criterios (a) resultados de un test de bandeja (*in basket*) y (b) puntuación obtenida en un cuestionario de evaluación del desempeño. En su estudio, se basaron en la taxonomía utilizada en el *Hogan Personality Inventory* (HPI; R. T. Hogan, 1986; R. T. Hogan y Hogan, 1992) y en el PRF (D. N. Jackson, 1984). La taxonomía correspondiente a HPI especifica de cuatro a ocho facetas pertenecientes a siete factores (que pueden ser clasificados según el modelo de los Cinco Grandes). A su vez, el PRF abarca 20 escalas específicas también y es posible obtener una estructura de cinco factores de la matriz de correlaciones de las mismas (estos factores son similares a los Cinco Grandes, aunque no exactos, Costa y McCrae, 1988). En ambos casos, los factores resultaron ser buenos predictores de los criterios, sin embargo, el análisis de la relación entre las

facetas de los factores y las facetas de los criterios aclaró cuáles de los rasgos son en realidad responsables de las relaciones existentes entre los factores y los criterios. Cabe mencionar que los coeficientes de correlación encontrados para las facetas tuvieron una validez promedio más pequeña que los coeficientes encontrados para los factores. Aunque en este caso, para comparar el poder predictivo de las facetas frente al de los factores se utilizó la correlación canónica y no la regresión múltiple, otra vez no se ha tenido en cuenta la triple composición de la varianza de las facetas. Además, es importante resaltar que cada uno de los instrumentos utilizados mide facetas diferentes los Cinco Grandes. Otra de las limitaciones es que las escalas del cuestionario PRF cargan de forma significativa en más de un factor a la vez, lo que indica que no son las facetas puras, sino las variables compuestas. Por todo lo anterior, puede concluirse que los resultados de Tett y colaboradores (2003) no aportan pruebas consistentes a favor del uso de las medidas específicas.

Otro estudio en esta línea fue llevado a cabo por Griffin y Hesketh (2004). Su objetivo fue explicar por qué la apertura a la experiencia ha sido considerada de forma consistente un predictor poco útil en los meta-análisis sobre la personalidad y el desempeño (p. ej., Barrick y Mount, 1991; Barrick et al., 2001; Salgado, 1997). Griffin y Hesketh supusieron que la razón pueden ser las diferentes relaciones que muestran sus componentes con el desempeño. Como medida del factor, en la primera muestra se utilizaron los 48 ítems que miden apertura en el NEO PI-R (Costa y McCrae, 1992b), y en las dos muestras siguientes, los 48 ítems provenientes de *International Personality Item Pool* (IPIP; Goldberg, 1999). A su vez, se utilizó una taxonomía de dos dimensiones de apertura a la experiencia (apertura a las experiencias externas e internas). Como criterio se utilizaron las valoraciones del rendimiento provenientes de los supervisores inmediatos. Los autores esperaban que la apertura a las experiencias externas se relacionase de forma positiva con el desempeño, mientras que la apertura a las experiencias internas se relacionase de forma negativa, sin embargo los resultados de este análisis no fueron significativos. Se confirmó que la apertura a la experiencia no es un constructo muy útil para predecir el desempeño general, pero la cuestión de si los factores o las facetas son mejores predictores no ha sido resuelta. No obstante, es importante destacar que la taxonomía utilizada en este estudio ha sido en cierto modo

arbitraria. Al mismo tiempo, tampoco se consideró la varianza común de las facetas que se debe al factor de apertura como una variable subyacente.

Roberts, Chernyshenko, Stara y Goldberg (2005) a su vez, al igual que Steward (1999), analizaron la validez del factor conciencia y de sus facetas para predecir varios criterios, entre ellos, la dedicación al trabajo. El mérito de estos autores es la creación de su propia taxonomía de facetas de conciencia, basándose en un análisis factorial de 36 escalas relacionadas con este factor y provenientes de siete diferentes y reconocidos inventarios de personalidad (algunos sin embargo no creados para reflejar el modelo de los Cinco Grandes). Hallaron que la estructura de conciencia puede ser expresada por seis facetas, que además tienen relaciones diferentes que el factor general con los criterios analizados. Aunque la diferencia entre el poder predictivo de las facetas y el del factor no fue muy grande, en cada caso la correlación múltiple de las seis facetas explicaba más varianza que el uso de la puntuación factorial únicamente. Cabe resaltar no obstante, que los criterios utilizados en el estudio no fueron relacionados directamente con el desempeño laboral, con la excepción de la dedicación al trabajo. Aún así, las conclusiones provenientes de este estudio deben de ser tratadas con cautela, ya que de la varianza total de las facetas no se restó la parte de varianza debida a su pertenencia al factor de conciencia y al error de medida antes de utilizar el método de regresión múltiple, sesgando de este modo los resultados obtenidos.

En otro estudio reciente, Warr, Bartram y Martin (2005) analizaron la relación entre los Cinco Grandes y el desempeño en tres diferentes puestos de vendedores (de coches, de aparatos electrónicos y de libros de puerta en puerta), aportando los datos de la validez para las facetas de factores conciencia y extraversión (dos por cada factor). Cabe destacar que las facetas utilizadas se calcularon como los compuestos de las escalas más específicas que incluía el cuestionario aplicado (Customer Contact Styles Questionnaire, CCSQ: SHL, 1997) y que los factores a su vez se calcularon como los compuestos de estas facetas, una conceptualización incorrecta, ya que el factor no es una simple suma de facetas sino una variable latente. Warr y colaboradores hallaron que conciencia es un buen predictor del desempeño de vendedores de coches y de libros (.26 y .21 respectivamente). A su vez, la relación entre la extraversión y el desempeño no fue significativa. Analizando los datos obtenidos concluyeron que el

poder predictivo de conciencia fue consecuencia de la relación positiva y significativa de la faceta de orientación al logro con el desempeño en todos los puestos (.26 para vendedores de coches, .30 para vendedores de aparatos electrónicos y .25 para vendedores de libros), mientras que la relación entre el desempeño y la faceta de confiabilidad no fue significativa en ningún puesto. Pero, esta conclusión no es correcta ya que al ser una variable latente, el factor tiene validez por si mismo y no debido a la relación que tienen sus facetas con el criterio. En realidad ocurre lo contrario, las facetas tienen validez gracias a su dependencia de un factor de orden superior. De forma similar, en el caso de extraversión, Warr y colaboradores otra vez concluyeron incorrectamente que su relación positiva con el desempeño se debió a la relación positiva de la faceta potencia con el desempeño en todos los puestos, aunque solo en el caso de los vendedores de coches, el coeficiente de validez fue significativo (.26), mientras que la faceta de afiliación tuvo una relación negativa, pequeña y no significativa con el desempeño. Además, para calcular el poder predictivo de las facetas en este estudio también se utilizó el método de regresión múltiple que en sí no resta a las facetas la varianza atribuida a los factores subyacentes y al error. Otra cuestión problemática es que el CCSQ: SHL no fue desarrollado dentro del marco de los Cinco Grandes, lo que aumenta el error de la medida. Los autores no han demostrado que las escalas de este cuestionario carguen cada una en un solo factor, lo que sería necesario para poder hablar de las *facetas*. Todas estas desventajas impiden la generalización de los resultados de Warr y colaboradores.

Christiansen (2008) también se mostró partidario de maximizar la predicción mediante el uso de escalas específicas. Comparó el poder predictivo de las facetas con el de los factores respecto a varias dimensiones del desempeño (efectividad interpersonal, toma de decisiones, supervisión, automotivación, e influencia) y al desempeño general, en una muestra de 262 mandos intermedios. Tanto las puntuaciones de los factores como de las facetas (4 por cada factor) provenían del *Global Personality Inventory* (GPI; Schmit, Kihm y Robie, 2000). A su vez, la varianza específica de las facetas fue tratada de tres formas diferentes, siguiendo distintos enfoques. Primero, se la consideró varianza de error, por lo que no se la ha tenido en cuenta, centrándose únicamente en la varianza explicada por el factor. Segundo, se consideró que las facetas aportan validez añadida a la varianza de los factores y que

esta varianza es igual de importante en cada caso, por lo que se calculó su validez predictiva creando un compuesto de las mismas (sin ponderarlas). Finalmente, se ponderaron las facetas según los pesos asignados a cada una en el método de regresión múltiple, prescindiendo de algunas facetas no relevantes para el criterio. Los resultados de estos análisis indican que mediante la primera estrategia se explicó el 40% de la varianza, menos que aplicando dos estrategias restantes, independientemente de si el criterio fueron las medidas específicas del desempeño o el desempeño general, e independientemente de la corrección de los resultados por la contracción (*shrinkage*). Sin embargo, la más efectiva de todas fue la estrategia de ponderación de las facetas relevantes para el criterio. Aunque Christiansen ve estos resultados como una confirmación de la superioridad predictiva de las facetas, su conclusión debe ser tratada con cautela debido a los errores metodológicos que posee su estudio. Cabe destacar que el autor consideró el factor una simple suma de facetas, mientras que, en realidad, es el factor el que influye en las facetas. Nuevamente, no se ha tenido en cuenta que la varianza de las facetas se compone de tres elementos, tratando el total de la varianza de las mismas como equivalente a su varianza específica. Del mismo modo, otro de los puntos débiles del estudio es que las facetas analizadas no formaban parte de una taxonomía reconocida, impidiendo realizar las inferencias y comparaciones con otros estudios similares.

El dilema amplitud-fidelidad reflejado en los estudios meta-analíticos

El primer estudio meta-análitico que analizó parcialmente la cuestión del dilema amplitud-fidelidad fue el de Hough (1992). Aunque la autora no usó el marco de los Cinco Grandes para resumir sus resultados, entre los predictores utilizados se encontraban los factores similares a los Cinco Grandes; entre ellos también dos reconocidas facetas del factor conciencia: logro y confiabilidad. Los resultados de este estudio indicaron que ambos constructos mostraban relaciones diferentes con los criterios utilizados. El logro predijo el desempeño general (.19), el comportamiento irresponsable (-.19), la efectividad en ventas (.27), trabajo en equipo (.14) y el esfuerzo (.21). La confiabilidad se relacionó a su vez con el comportamiento irresponsable (-.24) y con el trabajo en equipo (.17) y el esfuerzo (.14). A pesar de que Hough no trataba

estos dos constructos como componentes de un factor más amplio, los resultados encontrados llevan a la conclusión de que dependiendo del tipo de criterio que se quiera predecir, la elección de una u otra faceta de conciencia sería más conveniente.

En el caso del estudio de Hough (1992) también nos enfrentamos a una serie de problemas, muy similares a los encontrados en los estudios primarios sobre este tema. En realidad, no sabemos a qué amplitud del constructo nos estamos enfrentando, ya que Hough trata las escalas utilizadas como factores independientes. No podemos constatar con toda la seguridad que ambas facetas mencionadas dependan del mismo factor, ya que no disponemos de la matriz factorial de datos. En realidad se habla aquí de las facetas de conciencia teniendo en cuenta únicamente la similitud de terminología, pero es igualmente probable que tanto la escala de orientación al logro como la confiabilidad midan también otros factores. Por la misma razón, no se calculó cuánta varianza de los constructos analizados por Hough pertenece únicamente a éstos (varianza específica) y cuánta se debe a los factores del orden superior (comunalidad) y al error, por lo que los resultados obtenidos no aclaran las cuestiones relacionadas con el dilema amplitud-fidelidad.

Según Mount y Barrick (1995), la mayoría de los coeficientes de validez encontrados en los estudios meta-analíticos podría infravalorar el poder predictivo de los factores. Esto puede deberse al hecho de que la mayoría de estudios se basan en las escalas de personalidad individuales que provienen de los distintos cuestionarios de personalidad y que son medidas parciales e incompletas de los constructos de orden superior. En su estudio, Mount y Barrick se centran únicamente en el factor de conciencia. Señalan que en los meta-análisis anteriores se utilizó con frecuencia la validez promedio de las facetas de conciencia como un indicador de la validez de todo el constructo. Para evitar este error, los autores crean un compuesto de dos facetas de conciencia: logro y confiabilidad, utilizando la fórmula de Hunter y Schmidt (1990) para la corrección de puntuación de un compuesto (*composite score correction formula*)y para establecer la validez del factor de orden superior. A su vez, en la medición del criterio se basaron en una taxonomía multidimensional del desempeño que diferenciaba entre las medidas de desempeño generales y las medidas específicas. Las medidas de desempeño generales incluían la competencia (*proficiency*) laboral y la competencia en el entrenamiento. Las medidas específicas del desempeño abarcaban

competencia técnica, fiabilidad del empleado, esfuerzo, calidad, administración, orientación interpersonal, creatividad, y eficacia en situaciones de combate (*combat efficacy*). Aunque los autores señalan al principio que es posible que utilizando las facetas los criterios específicos se puedan predecir con más exactitud, los resultados de este estudio muestran que en cuanto a los criterios globales, existe muy poca diferencia entre la magnitud de los coeficientes de validez encontrados para conciencia, comparado con sus facetas. Se encontró un coeficiente de validez teórica para predecir el desempeño general de .31 para conciencia, y de .30 y .33 para confiabilidad y logro respectivamente. En cuanto a los criterios específicos, los autores concluyen que conciencia predice estos criterios igual o incluso ligeramente mejor que el logro y la confiabilidad (.40, .30, .38 respectivamente), contrario a lo que se podría esperar basándose en las premisas provenientes del dilema amplitud-fidelidad presentado por Cronbach y Gleser (1957). Los componentes de conciencia (confiabilidad y logro) predicen las medidas del desempeño específicas mejor que conciencia solo cuando están conceptualmente relacionadas con el criterio. Por ejemplo, confiabilidad fue mejor predictor de la fiabilidad del empleado (.47) y de la calidad (.48) que conciencia y logro. De la misma forma, logro era mejor predictor de esfuerzo (.58) y de creatividad (.19) que conciencia y confiabilidad. Sin embargo, la validez añadida de los componentes del nivel inferior respecto a conciencia era realmente pequeña y se encontraba entre .04 y .07.

A pesar de que las conclusiones de Mount y Barrick (1995) respecto a las propiedades predictivas de los factores son algo más optimistas que en los estudios anteriores, se han cometido similares errores metodológicos a la hora de diseñar el estudio y analizar los resultados. Primero, en ningún momento se estudió la estructura del factor de conciencia para elegir el número óptimo de las facetas que pertenecen al mismo. Segundo, los autores no aportaron los datos sobre la metodología utilizada para clasificar las escalas, por lo que no sabemos qué escalas han sido asignadas a cada una de las facetas y siguiendo qué criterio. Tercero, no se ha tenido en cuenta que el factor no es un simple compuesto de las facetas y que las facetas deben una parte de su varianza al mismo. Los autores asumieron de forma errónea que logro y confiabilidad son facetas “puras” de conciencia, por lo que analizando su validez predictiva trataron toda la varianza de estas facetas como equivalente a su varianza

específica. Estos errores analíticos hacen cuestionar los resultados y las conclusiones de este estudio meta-analítico.

El siguiente meta-análisis que abordó el tema de dilema amplitud-fidelidad fue el de Vinchur, Schippmann, Switzer III y Roth (1998). Los autores del mismo analizaron las propiedades predictivas de diferentes predictores del desempeño de vendedores y entre otros usaron los cinco grandes factores y las facetas de dos factores: extraversión (potencia y sociabilidad) y conciencia (logro y confiabilidad), siguiendo la clasificación de Hough (1992). Para evaluar el criterio se utilizaron dos tipos de medidas: subjetivas (las valoraciones de los superiores) y objetivas (el número de ventas). Los resultados indicaron que de los Cinco Grandes, solo extraversión y conciencia eran buenos predictores del desempeño de los vendedores, tanto de los criterios subjetivos (.18 y .21 respectivamente) como de los objetivos (.22 y .31 respectivamente). Sin embargo, las facetas de estos factores mostraron diferentes relaciones con el criterio. En el caso de ambos tipos de medidas del criterio, potencia resultó ser mejor predictor que extraversión (.28 y .26 respectivamente), al igual que logro resultó superar a conciencia en cuanto a su poder predictivo (.25 y .41 respectivamente). Al mismo tiempo, sociabilidad y confiabilidad resultaron tener un menor poder predictivo que los factores a los que corresponden: extraversión y conciencia, aunque el valor de sociabilidad no era generalizable. Estos resultados llevaron a los autores a aconsejar la medición en el nivel de facetas para analizar qué rasgos específicos son responsables de las relaciones significativas de los factores con los criterios, así como para maximizar la predicción.

Cabe destacar algunos puntos débiles de la metodología utilizada por Vinchur y colaboradores (1998). Primero, en ningún momento se estudió la estructura de factores conciencia y extraversión para saber el número óptimo de facetas para definir su dominio. Por ejemplo, en cuanto a conciencia, los estudios más recientes indican que este factor abarca por lo menos cuatro rasgos específicos (Dudley et al., 2006; Maulden, 2008; Roberts et al., 2005; Saucier y Ostendorf, 1999), por lo que una taxonomía de dos facetas únicamente puede resultar insuficiente. Segundo, al asignar las escalas provenientes de diferentes cuestionarios de personalidad a las facetas, aumenta la varianza de error. Esta varianza, junto con la varianza que se debe a la pertenencia de la faceta a un factor superior, explica una parte de la validez predictiva

de las facetas. Por lo tanto, para estimar la validez predictiva *real* de estas facetas habría que examinar únicamente su varianza específica, el procedimiento que no se ha llevado a cabo en este caso.

Otro meta-análisis o incluso una serie de los meta-análisis respecto a las cuestiones relacionadas con el dilema amplitud-fidelidad fueron llevados a cabo por Salgado (2003b). El objetivo de Salgado era comprobar si realmente las capacidades predictivas de las facetas son superiores a las de los factores para predecir los criterios específicos del desempeño: desempeño global de tarea y desempeño global contextual, según la clasificación de Borman y Motowidlo (1993), y tres facetas del último: dedicación al puesto (iniciativa conciente), apoyo personal y apoyo organizacional, según la clasificación de Borman, Penner, Allen y Motowidlo (2001). También se analizó si los factores superaban a las facetas como predictores cuando el criterio era el desempeño global. Las escalas provenientes de los inventarios de personalidad basados en el modelo de los Cinco Grandes fueron clasificadas según las indicaciones de sus autores, mientras que en el caso de los inventarios de personalidad no basados en los Cinco Grandes sus escalas fueron asignadas a los factores siguiendo las descripciones proporcionadas por Digman (1990), Hogan y Hogan (1992) y Salgado (1998a). A su vez, las escalas de todos los inventarios fueron asignadas a una taxonomía de 18 facetas siguiendo la propuesta de Saucier y Ostendorf (1999). De acuerdo a Saucier y Ostendorf, las facetas de cada uno de los cinco grandes factores serían las siguientes:

Neuroticismo – Estabilidad Emocional

- Irritabilidad
- Emocionalidad
- Inseguridad

Extraversión – Introversión

- Asertividad
- Sociabilidad
- Impulsividad
- Actividad

Apertura a la Experiencia – Cierre a la Experiencia

- Imaginación
- Intelecto
- Perceptividad

Amigabilidad – Rudeza

- Calidez
- Gentileza
- Generosidad
- Modestia

Conciencia – Irresponsabilidad

- Fiabilidad
- Orden
- Industriosidad
- Decisión

El número de estudios incluidos en cada meta-análisis varía dependiendo del predictor y del criterio. En el caso de los factores, el mínimo de estudios utilizados era 24 (para apertura como predictor del desempeño contextual) y el máximo 133 (para la conciencia como el predictor del desempeño global en el puesto). A su vez, el número de sujetos también fue diferente, siendo el mínimo 3,869 personas (para apertura como predictor del desempeño contextual) y el máximo 33,668 (para la conciencia como el predictor del desempeño global). En cuanto a las facetas, el mínimo de estudios analizados fue 3 y el máximo 19, siendo el número mínimo de sujetos incluidos en el análisis 368 (para la impulsividad y el desempeño global en el puesto), y el máximo 2,834 (para la sociabilidad y el desempeño global). En el caso de algunas facetas de los Cinco Grandes y algunas facetas del desempeño contextual, el número de estudios encontrados era demasiado pequeño y no todas las combinaciones teóricas entre ambos pudieron llevarse a cabo.

En el caso de la relación de las medidas globales de la personalidad y las medidas globales y específicas del criterio, Salgado (2003b) efectuó el meta-análisis

haciendo las correcciones por error de muestreo, fiabilidad del predictor, fiabilidad del criterio y restricción indirecta del rango. En el caso de las facetas, al no existir toda la información necesaria para realizar estas correcciones, tuvo en cuenta únicamente el error de muestreo, la fiabilidad del predictor y del criterio. Encontró que dos factores fueron predictores generalizables del desempeño de tarea y contextual: conciencia y estabilidad emocional. Los coeficientes de validez teórica (verdadera) correspondientes fueron .33 y .21. Los mismos factores resultaron ser predictores generalizables de dimensiones específicas del desempeño: desempeño de tarea (.35 y .25 respectivamente) y de desempeño contextual (.33 y .18 respectivamente). En cuanto al meta-análisis de los factores de personalidad y las facetas del desempeño contextual, los estudios encontrados permitieron analizar la validez únicamente para dos facetas: iniciativa inconsciente y apoyo personal. En el caso de la primera, solo conciencia resultó ser un predictor generalizable (.30), en cambio, en el caso de la segunda, tres factores resultaron ser predictores válidos: conciencia (.21), estabilidad emocional (.20) y amigabilidad (.24). En cuanto a las combinaciones entre diversas facetas de las dimensiones de personalidad y el desempeño global, 14 de las 18 facetas resultaron ser predictores válidos del mismo, aunque los tamaños de validez verdadera varían substancialmente de .05 para modestia (faceta de amigabilidad) hasta .21 para asertividad (faceta de extraversión) y fiabilidad (faceta de conciencia). En general, incluso aquellas facetas que mostraron los coeficientes de validez teórica más elevados, no superaron los coeficientes alcanzados por los factores conciencia y estabilidad emocional. Sin embargo, en cuanto a las facetas de los factores de extraversión y apertura, los resultados surgieren que desde el punto de vista aplicado, las facetas de asertividad y de la perceptividad deberían incluirse en los procesos de selección, ya que demuestran los coeficientes de validez para predecir el desempeño global superiores a los de los factores correspondientes. A su vez, los factores de conciencia y estabilidad emocional resultaron ser buenos predictores también de los criterios específicos, por lo que, en líneas generales, este estudio no apoya la hipótesis de aquellos autores que surgieren que las facetas pueden ser mejores predictores del desempeño global en algunas circunstancias.

El estudio de Salgado (2003b) posee varias ventajas, pero también presenta algunos inconvenientes. Entre las primeras cabe destacar que es el trabajo meta-

analítico más amplio sobre la cuestión del dilema amplitud-fidelidad y las medidas de personalidad realizado hasta la fecha. Además, es el primero que tanto para el predictor como y para el criterio utiliza taxonomías estudiadas y reconocidas: en el caso de las facetas de los Cinco Grandes la taxonomía de Saucier y Ostendorf (1999), y en el caso de las facetas del desempeño laboral la taxonomía de Borman y Motiwidlo (1993) y de Borman y colaboradores (2001). Sin embargo, entre los inconvenientes cabe destacar que tampoco se tuvo en cuenta la triple composición de la varianza de las facetas, por lo que del total de la varianza de las mismas no se descontó el valor de comunalidad y del error de medida, lo que indica que la validez de las facetas fue sobreestimada.

Otro estudio meta-analítico más reciente que analiza las cuestiones relevantes para el dilema amplitud-fidelidad es el de Dudley y colaboradores (2006). El objetivo de este meta-análisis fue analizar si las facetas de conciencia (en este caso cuatro: logro, confiabilidad, orden, y prudencia) tienen propiedades predictivas superiores al factor en cuanto a varios criterios del desempeño: desempeño general, desempeño de tarea, dedicación al trabajo, facilitación interpersonal y conductas contraproductivas. Los cuatro primeros fueron prestados por Hurtz y Donovan (2000), y el quinto fue prestado por Salgado (2002). Los coeficientes de la validez predictiva de conciencia correspondientes a los arriba mencionados criterios del desempeño fueron .24, .16, .20, .18 y -.26. Además, se distinguieron las siguientes categorías ocupacionales: vendedores, trabajadores de atención al cliente, gestores, trabajadores cualificados y semicualificados. También se consideraron las posibles variables moderadoras como el tamaño o la naturaleza de la muestra (los ocupantes del puesto frente a los candidatos). Los rasgos específicos alcanzaron coeficientes de validez predictiva para el desempeño de tarea desde .11 hasta .25, siendo el valor más alto el correspondiente al logro. Para la dedicación al trabajo, los coeficientes varían de .08 a .46, con el más alto correspondiente a la confiabilidad. Para la facilitación interpersonal, los coeficientes de validez predictiva variaron desde -.02 hasta .23, siendo igualmente el valor más alto el de confiabilidad. Finalmente, para las conductas contraproductivas los coeficientes variaron desde -.34 hasta .00 teniendo, una vez más, la confiabilidad alcanzando la validez más alta (el valor absoluto). Además, se analizó la validez añadida de las facetas en el caso de distintos criterios del desempeño. El porcentaje de varianza explicado

por los rasgos específicos por encima del factor de conciencia varía de 3.5% a 25.9%. Para el desempeño general, las facetas explicaron el menor porcentaje de la varianza adicional, solo 3.7%. En el caso del desempeño de tarea y la facilitación interpersonal, el porcentaje también fue pequeño, 4.6% y 5.8%, respectivamente. Para la dedicación al trabajo el porcentaje fue más significativo (25.9%). Finalmente, para las conductas contraproductivas la validez añadida alcanzó 13.6%. Los autores consideraron también la validez añadida de las facetas según los grupos ocupacionales. Para todas las ocupaciones, los resultados de la regresión indicaron un aumento significativo en la varianza explicada por uno o más rasgos específicos por encima del factor general de conciencia. El porcentaje de la varianza explicada varió del 1.2% al 24%, siendo el valor más bajo para los trabajadores del atención al cliente, y el más alto para los trabajadores cualificados y semicualificados. Para los puestos de gestión, el porcentaje alcanzó 9.3%, y para los vendedores 5.4%. Por lo tanto, los autores concluyeron que las facetas son predictores independientes de los factores y presentan un valor añadido incluso en el caso de los criterios globales.

Sin embargo, en este meta-análisis tampoco se evitaron los errores metodológicos similares a los encontrados en la mayoría de los estudios anteriores. A pesar de que en este caso se usó una taxonomía de cuatro facetas de conciencia, sobre la que más consenso existe entre los investigadores (J. Hogan y Ones, 1997; Saucier y Ostendorf, 1999) para calcular la validez añadida de los rasgos específicos, se aplicó el método de regresión múltiple. Tal y como se ha mencionado antes, este método en sí no tiene en cuenta que no toda la varianza que explican las facetas se debe a su varianza específica, y no descuenta la varianza perteneciente al factor general de conciencia y al error, por lo que las conclusiones provenientes del estudio de Dudley y colaboradores (2006) no pueden ser definitivas.

Un problema clave del debate: falta de una taxonomía de facetas

Un lector atento percibirá que es muy difícil sacar conclusiones generales de todos los estudios antes mencionados. Analizando la literatura existente al respecto, Rothstein y Jelly (2003, pág. 246) concluyeron: "No existe evidencia convincente sobre si las medidas de personalidad amplias o estrechas son preferibles para predecir el

desempeño laboral". Incluso teniendo en cuenta solo los meta-análisis al respecto, vemos que los resultados no siempre han sido coherentes. Asimismo, *todos* los estudios existentes, tanto los que apoyan el uso de las facetas, como los que apoyan el uso de los factores, poseen relevantes defectos metodológicos que impiden aceptar sus resultados y sacar conclusiones definitivas respecto al dilema amplitud-fidelidad.

Incluso sin tener en cuenta las deficiencias de los estudios antes mencionados, cualquier acuerdo sobre si las facetas o los factores son mejores predictores de los criterios organizacionales es poco probable mientras se usen taxonomías diferentes de facetas. Por ejemplo, Steward (1999) analizó las facetas de conciencia, logro y orden, en el caso de estudios de Hough (1992) y Mount y Barrick (1995), éstas fueron logro y confiabilidad. A su vez, Roberts y colaboradores (2005) enumeraron seis escalas que componen este factor, mientras que Dudley y colaboradores (2006) y Salgado (2003b) se basaron en una taxonomía de cuatro facetas. Recientemente, Maulden (2008) encontró que el modelo de conciencia de seis facetas es el que presenta el mejor ajuste a los datos, sin embargo, debido a alta correlación entre las escalas específicas, el modelo de 4 facetas podría resultar ser más práctico. De todas las facetas de conciencia, únicamente la faceta del orden fue enumerada en todos los estudios arriba mencionados, por lo que puede considerarse que el acuerdo sobre la misma es universal, por el contrario, en el caso de otras facetas de conciencia tal acuerdo no es unánime. A su vez, el consenso en cuanto a los rasgos específicos correspondientes a otros factores es aún menor. Aunque existen algunos estudios (p. ej., Dudley et al., 2006; Salgado, 2003b) que se basaron en una taxonomía de facetas de Saucier y Ostendorf (1999) reproducida en dos idiomas, alemán e inglés, siguen realizándose los estudios donde la elección de facetas es arbitraria.

Mientras que los investigadores no usen mayoritariamente una taxonomía de facetas ampliamente aceptada, el dilema amplitud-fidelidad será muy difícil de resolver. Incluso los cuestionarios de personalidad creados para medir el modelo de cinco grandes factores, a pesar de mostrar una validez convergente razonable, difieren respecto al número de las facetas asociadas a cada uno de los factores. Así, por ejemplo, el NEO PI-R (Costa y McCrae, 1992b) define 30 facetas, el HPI (R. T. Hogan y Hogan, 1992) 44 compuestos de ítems homogéneos, el instrumento AB5C (Hofstee et al., 1992) 45 facetas, y el *Global Personality Inventory* (GPI, Schmit et al., 2000) incluye

32 escalas específicas, el IP/5F (Salgado, 1996, 1998b) 29 Grupos de Ítems Homogéneos (Homogeneous Item Clusters – HICs) y el B5FI (Salgado, 2006) 18 facetas. Y si se añade que muchos de los estudios se basan en inventarios no desarrollados dentro del modelo de los cinco grandes factores, por ejemplo, 16PF (Cattell et al., 1970), el problema relacionado con la necesidad del uso de la única taxonomía de las facetas se hace aún más manifiesto.

Teniendo en cuenta la utilidad de una taxonomía de las subdimensiones de los Cinco Grandes, es sorprendente que casi no se hayan realizado estudios para evaluar la estructura de personalidad en el nivel menor de abstracción. Cabe destacar, sin embargo, el ya mencionado estudio de Saucier y Ostendorf (1999), cuyo objetivo era comprobar si existen facetas comunes de los Cinco Grandes en inglés y alemán, usando para ello los adjetivos marcadores de las cinco grandes dimensiones en ambos idiomas. Los autores han podido observar que en el marco de personalidad existe una estructura jerárquica de tres niveles. El primer nivel lo componen los rasgos de personalidad, asociados a cada uno de los adjetivos utilizados en el estudio. El segundo nivel lo componen las facetas o subdimensiones de la personalidad, y el tercer nivel lo componen las cinco grandes dimensiones de personalidad. Los resultados apoyan la existencia de 18 facetas reproducibles en ambos idiomas (tres para neuroticismo y apertura a la experiencia, y cuatro para cada uno de los factores restantes), y aunque existen diferencias en algunas facetas, hay también muchas similitudes entre ambos modelos.

Tradicionalmente, para estudiar la estructura de personalidad en todos los niveles de abstracción, los investigadores identificaban varios factores "primarios" oblicuos, que luego se colocaban en un modelo jerárquico mediante el análisis factorial del orden superior, basándose en las correlaciones entre estos factores primarios (estrategia "de abajo-hacia arriba", *bottom-up*). La novedad del estudio de Saucier y Ostendorf (1999) consiste en la utilización de la estrategia "de arriba-hacia abajo"*(top-down),* empezando con los factores amplios e independientes y buscando sus componentes a continuación. Esta estrategia parece más correcta considerando que el factor es una variable latente y no un compuesto de diferentes facetas.

Saucier y Ostendorf (1999) concluyeron que es posible que el 18 no sea el número final de las facetas reproducibles entre varios idiomas; al fin y al cabo, inglés y alemán son unos idiomas muy cercanos en cuanto a su génesis, lo que podría sesgar el resultado obtenido. Por consiguiente, recomiendan repetir el mismo estudio con otros idiomas provenientes de familias lingüísticas distintas, para confirmar o bien, modificar la estructura propuesta. Cabe destacar, sin embargo, que las medidas de personalidad utilizadas con mayor frecuencia comparten muchas de las facetas de la clasificación de Saucier y Ostendorf.

Por todo lo anterior, aunque los estudios que se basan en la estructura hallada por Saucier y Ostendorf son todavía muy escasos (p. ej., Dudley et al., 2006; Salgado, 2003b), esta taxonomía parece ser un paso muy prometedor hacia el descubrimiento de una taxonomía universal y aceptable de facetas.

Conclusiones generales del debate

Por el momento, a pesar del intenso debate, el dilema amplitud-fidelidad sigue sin resolverse. Las partes en el debate proporcionan argumentos convincentes, pero, a menudo contradictorios. Ya en los comienzos del debate, Briggs (1992) sugirió que el dilema amplitud-fidelidad es más complejo de lo que podría parecer, debido a que tanto el uso de los factores como de las facetas posee ventajas y desventajas. Cuando escalas amplias se relacionan con un criterio, generalmente no está claro por qué y cómo funcionan. Asimismo, las escalas específicas aportan claridad, sin embargo, limitan el rango de las aplicaciones. Por tanto, la amplitud aumenta la utilidad, pero a costa de la claridad conceptual, mientras que las medidas estrechas ofrecen claridad conceptual, pero a costa de la validez predictiva general de la escala.

Así, algunos autores que consideran que las propiedades predictivas de los factores superan a las propiedades predictivas de las facetas están a favor de la medición también a través de las escalas específicas. Según Chapman (2007), aunque los grandes factores permiten medir un constructo general y multidimensional con mayor exactitud, pueden apreciarse algunas diferencias significativas entre las facetas que los componen, debido a que cada relación encontrada entre el factor y el criterio es en realidad una combinación de posibles relaciones con las facetas. Por lo tanto, el

significado exacto de la relación entre el predictor y el criterio es difícil de interpretar en el caso de los factores, ya que no sabemos cuál de las facetas es responsable del tamaño de la relación. De este modo, la medición de las facetas ayuda a aclarar estos vínculos conceptuales entre dos constructos. Además, puede darse el caso que el factor no esté relacionado con el criterio en cuestión, pero sí alguna de sus facetas, entonces una mayor validez predictiva puede conseguirse utilizándola como un predictor independiente (Vasilopoulos, Cucina y Hunter, 2007).

A su vez, muchos de los investigadores que apoyan el uso de las facetas están de acuerdo en que la medición de los factores normalmente maximiza la predicción de los criterios globales (J. Hogan y Roberts, 1996; Moberg, 1998; Schneider et al., 1996). Como sugirió John y colaboradores (John, 1990; John et al., 2008; John y Srivastava, 1999), la personalidad puede ser conceptuada en distintos niveles de abstracción y amplitud, y cada uno de ellos puede hacer contribuciones únicas a nuestro entendimiento de la complejidad de la conducta humana. Sin embargo, para decidir cuál de los niveles de medición es más útil en un caso particular, se necesita más conocimiento acumulado (Hough y Furnham, 2003). Se necesita también conocimiento más objetivo, en el que las conclusiones se basen en los hechos y no en las opiniones personales de los autores. Por ejemplo, Rothstein y Goffin (2006, pág. 163) revisaron 11 estudios que tratan el dilema amplitud-fidelidad y concluyeron: "Juzgando en base a los estudios empíricos realizados desde el comienzo del debate, los rasgos específicos superan claramente las amplias dimensiones de personalidad". Sin embargo, una mirada más detallada a los estudios incluidos en esta revisión nos indica que: (a) en esta revisión se incluye por ejemplo, el estudio de Teet y colaboradores (2003), aunque los autores del mismo constatan que los resultados de su investigación no son concluyentes; (b) se incluyen los estudios de Conte y Gintoft (2005), de Crant (1995) y de Lounsbury, Gibson y Hamrick (2004), que tratan de los constructos que no están incluidos dentro de los Cinco Grandes, por lo que no podemos decir con certeza qué nivel de amplitud representan los rasgos analizados; (c) se incluye el estudio de Jenkins y Griffith (2004), que no usa el marco de los Cinco Grandes para resumir sus resultados, además de que los factores y las facetas evaluadas provienen de instrumentos diferentes; (d) entre los estudios revisados no se mencionan los estudios cuyas conclusiones han sido algo más favorable con la hipótesis de que las propiedades

predictivas de los factores son superiores que de las facetas, por ejemplo, el de Mount y Barrick (1995) o de Salgado y colaboradores (1995). Por todo lo anterior, cabe concluir que la revisión de Rothstein y Goffin presenta muchas lagunas que impiden extraer las conclusiones fiables de este trabajo.

Una conclusión de todo este debate es que analizando las propiedades predictivas de personalidad de las facetas, las investigaciones futuras deberían *utilizar una única taxonomía de facetas* o bien, *algún enfoque alternativo, que permitiese analizar el valor predictivo de las facetas sin centrarse en ninguna taxonomía en concreto.* Además, habría que *aplicar las técnicas analíticas correctas*. De este modo, en el nivel de facetas serán posibles las comparaciones de los resultados provenientes de estudios diferentes, tal y como hasta ahora se ha hecho en el nivel de los cinco grandes factores. Solamente de esta manera el dilema amplitud-fidelidad podrá ser finalmente resuelto.

Los objetivos del presente trabajo

Los objetivos de esta tesis son: (1) comprobar si los factores son mejores predictores que las facetas (o viceversa), y (2) si la amplitud del criterio a predecir influye en estas capacidades predictivas.

Para responder correctamente a la cuestión planteada, (a) las facetas serán identificadas mediante el análisis factorial y (b) se emplearán las técnicas analíticas adecuadas. Además, para responder de forma definitiva a las cuestiones relacionadas con el dilema amplitud-fidelidad, (c) se utilizarán los criterios de diferente amplitud, ya que varios autores (R. T. Hogan et al., 1996; Schneider et al., 1996) sugirieron que los factores serían mejores predictores de los criterios amplios y las facetas de los criterios más específicos.

Por el momento, en la mayoría de estudios realizados, se han utilizado taxonomías realizadas desde un punto de vista conceptual y no mediante un examen empírico. Por todo lo anterior, es imposible comparar los resultados obtenidos hasta el presente. A pesar de que la taxonomía de Saucier y Ostendorf (1999) parece ser un paso prometedor hacia una taxonomía de facetas ampliamente aceptada, todavía no tiene un apoyo mayoritario. Es por esto que, en la presente tesis, las facetas se

definirán mediante el análisis factorial. Esto nos permitirá tratar la varianza de las facetas de forma independiente de cualquier taxonomía que pueda utilizar un cuestionario de personalidad en concreto.

El segundo paso poder sacar conclusiones generalizables respecto al dilema amplitud-fidelidad consiste en la aplicación de las técnicas analíticas adecuadas. Tal y como se describió anteriormente, la mayoría de estudios relacionados con este dilema en el ámbito de personalidad aplicó el método de regresión múltiple sobre las correlaciones observadas, en vez de aplicarlo sobre las puntuaciones factoriales residualizadas de factores y de facetas. Además, en los estudios en los que se aplicaron métodos diferentes a la regresión múltiple, tampoco se ha tenido en cuenta que para calcular correctamente la validez predictiva de las facetas habría que separar primero de la varianza de las mismas esta parte que se debe al factor. Por todo lo anterior, en la literatura existente al respecto la validez predictiva de las facetas ha sido sobreestimada.

Para evitar las limitaciones de los estudios anteriores, en la presente tesis se aplicará un procedimiento que permite separar, en los cuestionarios de personalidad, la parte de la varianza que se debe a los Cinco Grandes de la que se debe únicamente a las facetas. En realidad, existen tres procedimientos diferentes que permiten realizar este análisis. Los dos primeros son *análisis de componentes principales (ACP)* y *análisis factorial (análisis de factores principales) (AF).* La principal diferencia entre estos métodos es que en el primero se usa toda la varianza de los ítems, mientras que en el segundo se usa solo la parte de varianza que cada ítem tiene en común con los otros ítems incluidos en el análisis. De este modo, mientras que el ACP busca formar las combinaciones lineales de las variables originales que expliquen la mayor parte de la varianza total, el AF pretende establecer un nuevo conjunto de variables no observables, menor en número que las variables originales, que exprese la mayor parte de la varianza común. En la mayoría de los casos el ACP y el AF proporcionan resultados similares, sin embargo, el uso del primero suele ser más frecuente, debido a que es más fácil de interpretar. El tercer procedimiento posible es el *análisis factorial jerárquico (AFJ).* Consiste en la identificación de conjuntos de ítems que se rotan de forma oblicua y en calcular la correlación entre estos factores para, finalmente, proceder a un segundo análisis factorial, esta vez ortogonal. El conjunto de los factores

ortogonales obtenido de este modo, permite dividir la varianza total explicada por el modelo en dos partes: (a) la que se debe a la varianza común de los ítems (factores de segundo orden) y (b) la que se debe a los conjuntos de variables muy similares en el análisis (factores de primer orden). Schmid y Leiman (1957) desarrollaron un método que permite llevar a cabo este análisis y que transforma una solución proveniente del análisis factorial oblicuo en una solución ortogonal, preservando la simplicidad de la primera, pero a la vez ordenando las variables de forma jerárquica. Sin embargo, aunque hemos encontrado algunos programas estadísticos que incluyen este procedimiento, ninguno de ellos permite calcular la puntuación factorial en los factores de segundo orden, lo que a su vez impide el uso de este método en la presente tesis. Por lo tanto, aunque el uso de este procedimiento podría proporcionarnos la ventaja de disponer de un número de facetas reducido y probablemente interpretable, al no disponer del software informático adecuado, finalmente hemos optado por el método de componentes principales, debido a su fácil interpretación y accesibilidad (es el método de extracción utilizado por defecto en el paquete estadístico SPSS).

En el análisis de componentes principales, los componentes se ordenan de manera decreciente, estando compuesto el primer factor por la combinación lineal de las variables que explican la mayor cantidad de varianza (Nunally, 1978). El segundo componente, lo constituye la combinación lineal de variables que explican la mayor proporción de varianza residual, siempre cumpliendo con la condición de que no se relacionen con la primera combinación de variables que forman el primer componente. Una vez obtenido el segundo componente, la variabilidad restante da lugar al tercer factor o componente, y así, seguidamente, hasta explicar la variabilidad total, siempre respetando la condición de ortogonalidad. De este modo, la variabilidad explicada por todos los componentes principales siempre es máxima. Además, los componentes conservan la variabilidad inicial, es decir, la suma de las varianzas de los componentes es igual a la suma de las varianzas de las variables originales, y la varianza generalizada de los componentes es igual a la original.

En el caso del análisis de componentes principales aplicado a los cuestionarios de personalidad, el primer factor resultante de la solución no rotada, esto es, el que explica la máxima cantidad de la varianza común de todas las escalas que pertenecen a

la misma dimensión de personalidad, corresponde lógicamente al nivel de los cinco grandes factores. A su vez, los restantes factores explican la varianza perteneciente a las facetas (la varianza específica). Aun así, como argumenta McDonald (1981), hay que tener en cuenta que unidimensionalidad empírica (verificada) no siempre implica unidimensionalidad conceptual, ya que el hecho de que siempre aparezca un primer gran factor general viene dado por las propiedades del método de componentes principales en sí (el primer factor siempre explica la mayor parte de la varianza). Si bien es cierto que la existencia del único factor importante antes de la rotación no demuestra la unidimensionalidad del resultado, hay que tener en cuenta que en el caso de las medidas de personalidad disponemos de suficiente base teórica como para apoyar este supuesto. En el caso del nuestro estudio, las escalas factorializadas provienen de tres cuestionarios de personalidad diferentes, creados específicamente para medir los Cinco Grandes, y cuya validez de constructo ha sido probada en las investigaciones anteriores. Por ello, cabe esperar encontrar un único factor latente (conciencia, estabilidad emocional, amigabilidad, apertura, o extraversión) que subyace a todas las escalas pertenecientes a la misma dimensión de personalidad, independientemente del número de inventarios utilizados.

Luego se calcularán las puntuaciones factoriales en cada uno de los factores no rotados para todos los sujetos en la muestra. Posteriormente, estos pesos residualizados ("puros") se aplicarán en el método de regresión múltiple, subsanando de este modo los defectos metodológicos de las investigaciones anteriores, donde incorrectamente se han utilizado las correlaciones observadas. De esta manera, estimaremos la validez predictiva correspondiente a cada nivel de medida y estaremos en condiciones de comprobar si las facetas de personalidad efectivamente añaden la varianza a los grandes factores.

Finalmente, para comprobar si la cuestión de amplitud del criterio influye en las capacidades predictivas de los factores y las facetas, se utilizarán varios criterios que pueden definirse como más o menos amplios. Los criterios a predecir serán relacionados con el desempeño académico, un concepto que está ganando importancia en el ámbito de psicología de trabajo (N. Schmitt et al., 2009).

Munson y Rubenstein (1992) ofrecieron una posible explicación al hecho del por qué los rasgos de personalidad deberían correlacionarse con el desempeño

académico, aun teniendo en cuenta que las medidas de personalidad no fueron creadas específicamente para medir este constructo. Sugirieron, que debido a que el desempeño laboral es un reflejo del desempeño académico, existen las relaciones entre la personalidad y el desempeño académico, porque se han demostrado las relaciones de la personalidad con el desempeño laboral (Barrick y Mount, 1991; Tett et al., 1991). Esta premisa de que ambos conceptos se relacionan, mantiene el interés de los psicólogos organizacionales en el criterio del desempeño académico hasta el presente. Ya en los años 80, varias revisiones cualitativas llegaron a la conclusión de que cuanto mayor sea el desempeño académico, más posibilidades de éxito laboral existen (Baird, 1985; Samson, Graue, Weinstein y Walberg, 1984). Más recientemente, los autores de dos revisiones meta-analíticas sobre el desempeño académico llegaron a conclusiones muy similares (Kuncel, Credé y Thomas, 2007; Kuncel, Hezlett y Ones, 2004). Concluyeron que las tareas académicas son complejas y vagamente definidas, que exigen múltiples habilidades para llevarlas a cabo con éxito, al igual que ocurre en el caso de las tareas laborales. A su vez, Trapmann, Hell, Hirn, y Schuler (2007) observaron que el éxito académico es muy similar al éxito en la formación como faceta del desempeño laboral y sugirieron que por esta razón, desempeño académico y laboral pueden estar relacionados. Con anterioridad, Roth, BeVier, Switzer y Schippmann (1996) demostraron que la relación entre ambos tipos del desempeño existe, si bien es moderada (.35). En definitiva, aunque no se puede constatar que el desempeño laboral y académico sean exactamente el mismo constructo, efectivamente existen similitudes entre ambos, lo que convierte desempeño académico en un constructo relevante desde el punto de vista de psicología de organizaciones.

En este trabajo se utilizarán tres medidas distintas del criterio del desempeño académico, que además tienen amplitud diferente:

(1) *Las calificaciones académicas*, que de forma más completa reflejan el esfuerzo del alumno para superar exitosamente los exámenes u otras exigencias formales de la carrera. En este sentido, la nota media en el ámbito educativo se aproxima a las valoraciones del desempeño en el ámbito de las empresas.

(2) *Las conductas desviadas académicas.* Varios autores sugirieron que las conductas desviadas académicas puede estar vinculadas con la contraproductividad laboral. Blankenship y Whitley (2000) descubrieron que las conductas desviadas académicas tales como copiar durante los exámenes o evitarlos utilizando excusas falsas, se relacionan con muchas conductas desviadas generales. Lawson (2004) halló que mostrar conductas desviadas académicas (copiar o plagiar los trabajos académicos), implica menores puntuaciones en la escala de creencias éticas relacionadas con el mundo laboral. Lucas y Friedrich (2005) afirmaron que si se conceptúa el rol de *estudiante* como un tipo de *trabajo*, las conductas desviadas en el ámbito académico podrían entenderse como un tipo de conductas organizacionales desviadas. En este sentido serían una faceta del desempeño académico general.

(3) *Las medidas específicas (narrow)* (*General Competency Tests*). Este tipo de medidas se basan en una conducta puntual y miden desempeño académico en una situación concreta, por lo que son el criterio más específico de los tres utilizados.

Así pues, tenemos tres tipos de criterios con medidas independientes, y se utilizan diferentes métodos (calificaciones, conductas y tests), de los que unos son más amplios y otros más específicos. Todo esto permitirá cumplir con el objetivo de poder sacar conclusiones sobre si los factores superan a las facetas como predictores, y de averiguar si las propiedades predictivas de ambos varían dependiendo de la amplitud del criterio a predecir.

Personalidad y éxito académico – resumen de la investigación e hipótesis

El primer estudio que centró la atención de la comunidad científica en las variables no cognitivas, en concreto, en una serie de rasgos temperamentales como variables que pueden estar relacionadas con el éxito académico, fue llevado a cabo por Webb en 1915. En él, Webb encontró un factor que denominó "persistencia de motivaciones" (pág. 60) definido como "*consistencia de acción que resulta de la intención deliberada, es decir, la voluntad*" (pág. 60). El autor le asignó una abreviatura *w (factor w).* Las cualidades que tenían las cargas factoriales más altas en este factor fueron, por ejemplo: *tendencia de no abandonar la tarea únicamente para variar*, *tendencia de no abandonar la tarea ante dificultades*, *ser digno de confianza*, *ser concienzudo*. Las características del factor *w* surgirían de que éste pudiese estar relacionado con el rendimiento escolar y académico; efectivamente, Webb encontró que las cualidades que más alto cargaron en este factor también estaban altamente relacionadas con el rendimiento académico. Fue un hallazgo muy novedoso ya que, hasta aquel entonces, el éxito académico era considerado únicamente un reflejo de las capacidades intelectuales de las personas. Sin embargo, a partir de este momento, muchos investigadores empezaron a incluir las variables no cognitivas en los estudios dedicados al éxito o fracaso educacional. Las investigaciones posteriores confirmaban la importancia de factor *w* para el rendimiento académico y seguían añadiendo más constructos de personalidad a la lista de características cruciales para el éxito académico, entre otros: cooperación, integridad, dinamismo (*drive*), interés, motivación, curiosidad, creatividad y valores (Alexander, 1935; McCloy, 1936; Messick, 1979; Wechsler, 1943).

Al mismo tiempo, aparecieron varias revisiones cualitativas de la literatura al respecto (Harris, 1940; Margrain, 1978; Stein, 1963). Sin embargo, hasta que no hubo una denominación común para resumir todos los estudios existentes, las conclusiones sobre la relación entre los rasgos de personalidad y el éxito académico fueron confusas. Esta denominación común se hizo posible con el consenso respecto a los Cinco Grandes como una taxonomía de los rasgos de personalidad (Digman, 1990;

John, 1990), y con el desarrollo del meta-análisis como un método de acumulación de conocimiento (Hunter y Schmidt, 1990). A partir de este momento empezaron a aparecer algunos estudios que intentaron resumir toda la investigación existente de forma cuantitativa. Cabe destacar que a pesar del elevado interés en la relación entre el desempeño académico y la personalidad, el número de los estudios meta-analíticos al respecto no es elevado, en concreto, si lo comparamos con el número de meta-análisis dedicados a la relación de personalidad con el desempeño laboral.

A continuación, se presentan las conclusiones provenientes de los meta-análisis existentes y de la mayor revisión cualitativa dedicadas a las relaciones entre la personalidad y el éxito o el fracaso académico en el nivel de los factores. También se revisan los estudios más importantes que analizan estas relaciones en el nivel de las facetas. Posteriormente, se presentan nuestras hipótesis basándose en los resultados de las investigaciones más representativas en este ámbito.

Los factores como predictores de las calificaciones académicas

El primer estudio meta-analítico que incluye los rasgos de personalidad relacionándolos, entre otros, con un criterio de éxito académico es el estudio de Hough, Eaton, Dunnette, Kamp y McCloy (1990), cuya metodología fue descrita en el Capítulo I del presente trabajo. Los resultados del mismo indican que Dominancia/Energía (*Surgency*), que corresponde al factor de Extraversión, se relaciona moderadamente con éxito académico (.15), Adaptación (*Adjustment*), que corresponde al factor de Estabilidad Emocional, muestra la relación de .26, Confiabilidad (*Dependability*) y Achievement (Logro), que son facetas de Conciencia (aunque los autores no las tratan como las facetas, sino como los constructos independientes, clasificándolos, de hecho, en *categorías* diferentes), muestran la relación de .15 y .30, respectivamente. Finalmente, Intelectualidad/Cultura (*Intellectance*), que corresponde al factor de Apertura a la experiencia, alcanzó el coeficiente de correlación de .18. Estos coeficientes no fueron corregidos por la falta de fiabilidad, ni tampoco por la restricción del rango, lo que indicaría relaciones aun mayores si se llevasen a cabo estas correcciones. Sin embargo, al no presentar los

intervalos de confianza correspondientes, no sabemos cuáles de estas relaciones serían generalizables. Además, cabe destacar que al presente estudio pueden dirigirse las mismas críticas que a muchos de los estudios mencionados con anterioridad: que los constructos de personalidad utilizados no corresponden exactamente al modelo de los Cinco Grandes y que las escalas que formaban parte de los constructos de personalidad (con la excepción de la escala heterogénea) tenían relativamente moderadas correlaciones entre si: de .33 a .46 (con una media de .39). En último lugar, es importante destacar que el número de estudios incluidos en este meta-análisis era muy desigual para cada uno de los factores, y varía de 44 estudios para adaptación a 6 estudios en el caso de intelectualidad. Asimismo, los autores no indicaron el tamaño de dichas muestras, por lo que es imposible comparar sus tamaños. Por todo ello, sobre el estudio de Hough y colaboradores podemos decir únicamente que demuestra que existe cierta relación entre la personalidad y el éxito académico, aunque su magnitud real permanece desconocida, al igual que permanecen desconocidas las posibilidades de generalización de los resultados hallados. Esta confusión respecto a los vínculos entre los rasgos de personalidad y desempeño en el ámbito educativo se refleja también en la revisión cualitativa de De Raad y Schouwenburg (1996).

La revisión de la literatura sobre la personalidad y el éxito académico realizada por De Raad y Schouwenburg (1996) es la más amplia que se ha hecho hasta hoy, además de ser la primera que hace referencia al modelo de los Cinco Grandes. Los autores concluyeron que, efectivamente, la personalidad juega un papel importante en el ámbito educativo, en concreto, el factor de conciencia. Los autores concluyeron también que algunos rasgos relacionados con los factores de extraversión, estabilidad emocional y cultura (apertura a la experiencia) podrían estar relacionadas con el desempeño académico. Cabe destacar, sin embargo, que al tratarse de una revisión cualitativa, no se indicó el tamaño de ninguna de estas relaciones.

El siguiente estudio relevante al respecto es el de Salgado (2000). Salgado incluyó el éxito académico como uno de los criterios en su meta-análisis sobre los Cinco Grandes y los criterios alternativos del desempeño, al lado de conductas organizacionales tales como: absentismo, accidentes, compensación, contraproductividad, ventas y rotación, entre otros. Para los cálculos se utilizaron los

estudios relevantes sobre el tema llevados a cabo desde 1990 hasta 1999, y realizados en Europa y en Estados Unidos. En total utilizó 16 estudios, aunque el número de estudios y de sujetos varía dependiendo del factor. Y así, para calcular los coeficientes de la validez predictiva para el factor de la estabilidad emocional empleó 14 estudios con una muestra total de 3,784 personas. Para la extraversión el número fue 12 estudios y 3,585 personas, para la apertura a la experiencia, 11 estudios y 1, 654 personas, para amigabilidad, 11 estudios y 1,654 personas y para conciencia, 15 estudios y 2,214 personas. La única dimensión que resultó predecir el desempeño académico con una validez generalizable era conciencia .28. Aunque el autor suponía que también estabilidad emocional y apertura predecirían este criterio, estas hipótesis no se confirmaron. El estudio de Salgado fue el primer estudio meta-analítico que, explícitamente, utilizó el marco de los Cinco Grandes factores para resumir los resultados respecto a la relación de personalidad y el desempeño académico. Las escalas de los inventarios no desarrollados dentro del modelo de los Cinco Grandes fueron clasificadas según el mismo por dos jueces. Además, los resultados de esta revisión cuantitativa fueron corregidos por falta de fiabilidad en el criterio y en el predictor, al igual que por la restricción del rango. Se indicaron también los intervalos de confianza, lo que permite sacar conclusiones sobre la generalización. Entre las limitaciones del estudio cabe destacar el moderado tamaño de las muestras.

Unos años más tarde, Poropat (2005) llevó a cabo el siguiente resumen cuantitativo de la relación entre la personalidad y el desempeño académico. En él, incluyó todos los estudios publicados desde 1990 a 2004 llevados a cabo en muestras de los estudiantes universitarios, y únicamente con instrumentos basados en el modelo de los Cinco Grandes. Además, incluyó solamente los estudios que utilizaban como criterio las calificaciones o medias académicas. La muestra total abarcó 16 estudios y 4,626 estudiantes universitarios. Los resultados de este análisis indicaron que aunque todos los coeficientes encontrados fueron generalizables, el valor más alto correspondía a conciencia (.21).

Cabe observar, sin embargo, que Poropat (2005) utilizó en su estudio la correlación promedio ponderada por el tamaño muestral, sin aplicar la corrección por la falta de fiabilidad en el predictor y en el criterio o por la restricción del rango. Además, al basarse únicamente en los estudios publicados, los resultados de este

meta-análisis son susceptibles al sesgo de publicación selectiva que consiste en que solo se publican los estudios con los resultados positivos, mientras que los estudios que no encuentren relaciones significativas no suelen publicarse (Rosenthal y DiMatteo, 2001). Para terminar, aunque el criterio utilizado fueron las calificaciones, en la práctica se incluyeron las medidas muy variadas de las mismas, tales como la media académica proveniente de los registros oficiales, la auto-valoración de la media académica, la calificación correspondiente a una asignatura en particular o la media académica de un curso en concreto. Aún así, y a pesar de que Poropat no utilizó las correcciones meta-analíticas propuestas por Hunter y Schmidt (1990), cabe destacar como un punto a favor, que se utilizaron únicamente las medidas basadas en los Cinco Grandes, y que el desempeño académico siempre fue medido como calificaciones (incluso teniendo en cuenta la variabilidad de los tipos de las mismas). Otro punto a favor de esta revisión, es que fue realizada en el nivel universitario únicamente. Recapitulando, el estudio de Poropat volvió a confirmar el hallazgo previo de Salgado que de los Cinco Grandes, el factor de conciencia, exclusivamente, predice el desempeño académico, con una validez moderada.

Relativamente poco tiempo después del estudio de Poropat (2005) se llevó a cabo un nuevo meta-análisis de las propiedades predictivas de personalidad respecto al desempeño académico. Trapmann, Hell, Hirn y Schuler (2007) meta-analizaron todos los estudios publicados sobre este tema entre 1980 y 2004 que incluían las muestras universitarias y utilizaban por lo menos una escala que correspondiese o pudiese ser asignada al modelo de los Cinco Grandes. Además, como novedad, el análisis se realizó respecto a tres criterios del éxito académico independientes: calificaciones, abandono universitario y satisfacción. En total se encontraron 58 estudios provenientes de 15 países diferentes que proporcionaron por lo menos 30 correlaciones por factor en cuanto al criterio de calificaciones académicas. En cuanto a los criterios restantes: abandono universitario y satisfacción, el meta-análisis se llevó a cabo solo en el caso de los factores que superaban tres correlaciones encontradas. Así pues, las muestras fueron repartidas como sigue. Para el criterio de calificaciones se encontraron 59 estudios que acumularon en total una muestra de 14,653 personas para neuroticismo, 54 estudios y 12,424 personas para extraversión, 41 estudios y 14,942 personas para apertura, 34 estudios y 12,452 personas para amigabilidad, 41 estudios y 10,855

personas para conciencia. Para el criterio de abandono escolar, los análisis se hicieron para el factor neuroticismo con 4 muestras y 2,714 sujetos, factor extraversión con 5 estudios y 3,916 sujetos, y factor apertura con 4 estudios y 2,194 sujetos. Finalmente, en cuanto a satisfacción, el meta-análisis fue posible únicamente para los factores de neuroticismo y extraversión, el primero con 8 estudios y una muestra total de 1,512 personas y el segundo con 4 estudios y la muestra de 660 personas. Los resultados obtenidos fueron corregidos por el error muestral y por la falta de fiabilidad en el criterio y en el predictor. No se aplicó la corrección por la restricción de rango, ya que como argumentan los autores, en este caso la restricción del rango no tiene lugar debido a que la selección de los alumnos no se realizó mediante las medidas de personalidad. Este argumento, sin embargo, no es del todo correcto, ya que incluso en los grupos aparentemente homogéneos puede haber restricción de rango en las puntuaciones de los rasgos de personalidad que varía entre el dos y el nueve por cien, tal y como lo demostraron Ones y Viswesvaran (2003). Además, es posible calcular esta restricción del rango (p. ej., Salgado, 1997; 2000), utilizando la desviación típica encontrada en la muestra como restringida, y la desviación típica presentada en los manuales de los inventarios de personalidad como no restringida.

Por lo que se refiere al criterio de calificaciones académicas, Trapmann y colaboradores (2007) supusieron que todos los factores de personalidad mostrarían una relación positiva con el mismo. Los resultados de este meta-análisis indicaron que solo factor de conciencia mostró relación significativa y generalizable con el desempeño académico (.27). El valor de apertura resultó ser positivo (.13), pero el intervalo de confianza incluía el cero, por lo que no era generalizable. Respecto al criterio de abandono académico, las hipótesis fueron que mayores puntuaciones en conciencia estarían ligadas a menor abandono universitario, mientras que mayores puntuaciones en apertura serían relacionadas con su aumento. Sin embargo, en este caso ninguna de las hipótesis pudo ser investigada. Para el factor de conciencia los autores no encontraron un número suficiente de estudios para llevar a cabo los análisis y para la apertura el intervalo de confianza incluía el cero. Lo mismo ocurrió con neuroticismo y extraversión, los otros factores para los que se efectuaron los cálculos. En cuanto al último criterio, satisfacción, se tuvieron en cuenta únicamente los factores de neuroticismo y extraversión. Las hipótesis iniciales fueron que la

estabilidad emocional mostraría una relación generalizable y positiva con el desempeño académico, y que extraversión se relacionaría de forma positiva con este criterio, pero que los coeficientes de validez serían mucho más pequeños. En el caso de neuroticismo, que mostró una relación de -.37 (generalizable), la hipótesis fue confirmada. En el caso de extraversión, el coeficiente obtenido era de .10, sin embargo, el intervalo de confianza incluía el cero. Además, se analizaron posibles variables moderadoras de la relación entre el desempeño académico y la personalidad como, por ejemplo: carrera estudiada, cultura, edad, fuente de publicación, etc. Gracias a este análisis, se probó que conciencia es no solo el único factor generalizable en cuanto a la predicción de las calificaciones académicas, sino que también su validez permanece prácticamente invariable para todas las variables moderadoras utilizadas. Además, se encontró que aunque el poder predictivo general de factor extraversión en cuanto a las calificaciones no es significativo, si lo es en el caso de algunos continentes, además de que sus relaciones con el desempeño académico varían considerablemente. Por ejemplo, en el este de Asia ser extravertido conlleva un significativo aumento de las calificaciones académicas (.36), mientras que en Australia les influye de forma negativa (-.09). Recapitulando, el estudio de Trapmann y colaboradores indicó que el único factor de personalidad que predice el desempeño académico utilizando las calificaciones académicas como el criterio es factor conciencia. También se halló que la estabilidad emocional es un buen predictor de la satisfacción. Además, se demostró que extraversión puede ser mejor o peor predictor del desempeño académico, dependiendo de la cultura dominante de cada país o continente.

A pesar de su indudable aportación al conocimiento sobre la relación entre la personalidad y desempeño académico, cabe destacar que el estudio de Trapmann y colaboradores (2007) muestra también algunas limitaciones que pueden influir en la generalización de sus resultados. Primero, algunos de los estudios meta-analizados estaban basados en inventarios no desarrollados según los Cinco Grandes, como las primeras versiones del 16PF (Cattell et al., 1970), o el CPI (Gough, 1957). Además, no se aplicó la corrección por la medición imperfecta del constructo propuesta por Hunter y Schmidt (1990), que podría evitar un posible sesgo en los resultados en este caso. Segundo, el meta-análisis se basó únicamente en los estudios publicados, lo que puede llevar a un sesgo debido a la publicación selecta (Rosenthal y DiMatteo, 2001). Aun así,

este meta-análisis fue el siguiente granito de arena para otra vez constatar que la dimensión de personalidad que más influye en el desempeño académico es conciencia.

El mismo año que el estudio de Trapmann y colaboradores (2007), fue publicado otro meta-análisis por O'Connor y Paunonen (2007). En comparación con el primero, los estudios incluidos abarcaron un periodo de tiempo más corto, desde 1990 hasta 2006, por lo que las muestras acumuladas para cada factor también fueron menores. Otra diferencia fue que se utilizaron únicamente los estudios basados en los instrumentos desarrollados según los Cinco Grandes. En total, O'Connor y Paunonen encontraron 31 muestras provenientes de 22 estudios diferentes. En la mayoría de los casos, el criterio utilizado fueron las calificaciones académicas, medias o calificaciones de una materia o un examen, pero en algunos casos se utilizó un criterio diferente: las valoraciones del desempeño en el aula, las valoraciones del desempeño en pruebas escritas, la nota del proyecto final, la nota de los ensayos, etc. Las muestras totales por factor fueron: 21 estudios y 5,091 sujetos para neuroticismo, 22 estudios y 5,161 sujetos para extraversión, 23 estudios y 5,878 sujetos para apertura y conciencia, y 19 estudios y 4,490 sujetos para amigabilidad. Los resultados muestran que solo en el caso de amigabilidad y conciencia se encuentra una relación generalizable con el desempeño académico, aunque el coeficiente de validez predictiva de amigabilidad es casi nulo. Otra vez, conciencia resultó ser el único factor con una relación significativa y generalizable con el desempeño académico, mostrando una relación de .24 corregida por falta de fiabilidad en el predictor. Como novedad, los autores de este meta-análisis intentaron realizar el análisis también en el nivel de las facetas de los Cinco Grandes, sin embargo, como constatan muy pocos estudios, proporcionaron los datos al respecto por lo que dicho análisis no fue posible. En vez de la misma, llevaron a cabo una revisión cualitativa de la que sacan conclusiones que en algunos casos, los rasgos de personalidad más estrechos que los Cinco Grandes pueden tener poder predictivo superior a los factores en cuanto al criterio del desempeño académico. No obstante, debido a que ninguno de los estudios mencionados descontó de la varianza de las facetas la varianza común que estas deben a la pertenencia a un factor del orden superior, las conclusiones sacadas por O'Connor y Paunonen no pueden considerarse apoyadas empíricamente.

Cabe destacar que el meta-análisis de O'Connor y Paunonen (2007) también tiene ciertas limitaciones. Una de ellas es que entre las medidas del desempeño académico utilizadas se encontraron los criterios más variados que las valoraciones en forma de calificaciones de los exámenes de conocimiento. Esto podría no suponer un grave problema, pero en algunos casos los autores tratan como resultados independientes los coeficientes de validez predictiva alcanzados tanto para la nota final, como para los componentes de la misma. Por ejemplo, en un estudio de Rothstein, Paunonen, Rush y King (1994) incluido en este meta-análisis se utilizaron tres criterios: desempeño en las pruebas escritas, en el aula y la nota media. O'Connor y Paunonen acumulan estos datos como provenientes de tres muestras independientes, mientras que el análisis del artículo original indica que el desempeño en las pruebas escritas y en el aula formaban parte de la nota final, tratándose de la misma muestra de alumnos. También en el estudio de Chamorro-Premuzic y Furnham (2003a) la nota final y la nota del proyecto final son medidas tomadas en la misma muestra de estudiantes. Además, otra limitación del presente meta-análisis, es que no se aplicó la corrección por falta de fiabilidad en el criterio. Cabe observar que esto, igualmente, pudo haber sido imposible al utilizar los criterios tan variados bajo única denominación del desempeño académico. Finalmente, en esta revisión cuantitativa se incluyeron únicamente estudios publicados, lo que pudo haber aumentado el tamaño de los coeficientes de correlación encontrados debido al sesgo de la publicación selectiva (Rosenthal y DiMatteo, 2001).

La revisión de los estudios meta-analíticos en este ámbito debe terminar con el trabajo de Poropat (2009). Es el meta-análisis más reciente y a su vez más completo llevado a cabo hasta la actualidad. En este estudio se incluyeron todos los estudios, tanto publicados como y no publicados hasta el año 2007, incluyendo todos los estudios de los meta-análisis anteriores en los que se utilizaron los instrumentos basados en el modelo de los Cinco Grandes. Además, como novedad, este meta-análisis se realizó de forma jerárquica, es decir, los análisis se llevaron por separado para la enseñanza primaria, secundaria y superior, lo que adicionalmente permite la comparación del poder predictivo de personalidad en todos los niveles educativos. Como criterio, se utilizaron únicamente las calificaciones (de exámenes, asignaturas, cursos completos) y las medias académicas, tanto provenientes de los registros

oficiales, como y auto-informadas. La muestra total abarcó 63 artículos publicados y 17 trabajos no publicados. Al factor de amigabilidad le correspondían 109 coeficientes y el total de 58,522 personas, 138 coeficientes y 70,926 personas correspondían a conciencia, 114 coeficientes y 59,554 personas a estabilidad emocional, 113 coeficientes y 59,986 personas a extraversión y 113 coeficientes y 60,442 personas correspondían a apertura. Cabe destacar que las muestras acumuladas en este estudio son incluso superiores al meta-análisis de segundo nivel de Barrick y colaboradores (2001).

Los resultados indicaron que apertura a la experiencia mostró una pequeña relación con el desempeño académico (.12) y conciencia mostró una relación moderada (.24). Las relaciones de los restantes factores de personalidad con el desempeño académico eran próximos o iguales a cero: amigabilidad .07, estabilidad emocional .02, extraversión .00. Los coeficientes de validez encontrados por Poropat (2009) fueron corregidos por la falta de fiabilidad en el predictor y en el criterio, aplicando diferentes valores de fiabilidad para la nota media académica, para la media académica auto-informada, al igual que para las calificaciones de los exámenes puntuales. Cabe recordar que no se aplicó la corrección por restricción del rango, debido a la falta de información sobre los ratios de acceso a la universidad. Con anterioridad se ha explicado, no obstante, que hay otros métodos que permiten la estimación de la restricción del rango. De este modo, si corregimos estos coeficientes por la restricción del rango encontrada en el meta-análisis de Salgado (2000), sus valores serían respectivamente: .16 para apertura a la experiencia, .29 para conciencia, .09 para amigabilidad, .02 para estabilidad emocional, y en el caso de extraversión .00.

Poropat (2009) encontró además, que la relación de todos los rasgos de personalidad con el desempeño en el nivel universitario disminuye cuanto más alto es el nivel educativo para todos los factores, con la excepción de conciencia. Este hecho revoca el argumento de Chamorro-Premuzic y colaboradores (2006) que sugirió que cuanto mayor es el nivel educativo, más importante es el papel de los rasgos de personalidad para predecir el éxito. Solo en el caso de amigabilidad se encontró el valor ligeramente superior para los alumnos de universidad que para los de secundaria .06 frente a .05. La relación de amigabilidad con el éxito académico bajó sin embargo de .30 en el nivel de primaria a únicamente .06 en el nivel superior. Algo parecido

ocurrió con apertura, cuyo poder predictivo en el caso del desempeño académico de los alumnos de primaria era .24, de secundaria .12 y en la universidad solo .07. Por todo lo anterior, puede concluirse que conciencia es el único factor de personalidad cuya influencia en el desempeño académico se mantiene en todos los niveles educativos de forma prácticamente invariable, además sin influir en la misma la edad de los alumnos universitarios, como demostraron los análisis adicionales. Los resultados hallados para todos los factores en el nivel de educación superior fueron: .06 amigabilidad, .23 conciencia, -.01 estabilidad emocional, -.01 extraversión y .07 apertura a la experiencia, que si se corrigen por la restricción del rango encontrada por Salgado (2000) alcanzan respectivamente: .08, .28, .-.01, -.01, y .09. En este nivel educativo, conciencia otra vez resultó ser el único factor con un valor de validez moderado. No obstante, una limitación muy importante de este estudio es que no se conoce el valor del intervalo de confianza, por lo que no se pueden sacar las conclusiones respecto a la generalización de los resultados encontrados. Tampoco se han incluido los datos necesarios para calcularlo. Teniendo en cuenta que este meta-análisis de Poropat es el trabajo más amplio que se ha publicado en este ámbito hasta el momento y que sus muestras totales superan en tamaño incluso las muestras del mayor meta-análisis sobre la relación entre la personalidad y el desempeño, el de Barrick y colaboradores (2001), es razonable suponer que dichos resultados serían generalizables, si se hubiesen llevado a cabo las estimaciones oportunas. Cabe recordar que en un meta-análisis anterior del mismo autor (Poropat, 2005), incluso siendo las muestras acumuladas más pequeñas, todos los coeficientes de correlación encontrados resultaron ser generalizables. Aún así, las conclusiones al respecto deben de ser cautas.

El meta-análisis de Poropat (2009) analizó también la relación entre los rasgos de personalidad e inteligencia para comprobar si existe una validez añadida de personalidad sobre el criterio que se considera el mejor predictor tanto del desempeño laboral (Bertua, Anderson y Salgado, 2005; Hunter y Hunter, 1984; Levine, Spector, Menon, Narayanan y Cannon-Bowers, 1996; Salgado y Anderson, 2002, 2003; Salgado, Anderson, Moscoso, Bertua y de Fruyt, 2003), como del desempeño académico (Strenze, 2007). Cabe resaltar que la relación de inteligencia con el desempeño académico no suscita dudas, debido a que las medidas de inteligencia se

desarrollaron con el propósito de predecir las diferencias individuales en el aprendizaje (Brody, 2000). Poropat encontró que la relación de inteligencia con el desempeño académico fue de .25, es decir menor que las encontradas por Kobrin, Patterson, Shaw, Mattern y Barbuti (2008) para los alumnos de secundaria (.28) y universitarios (.35), cuando éstas no fueron corregidas por la restricción del rango, y mucho menor que la de Strenze (.56), que fue basada en las muestras que tenían la restricción del rango pequeña o ninguna. De los Cinco Grandes, el factor cuya relación con la inteligencia fue mayor resultó ser apertura, lo que, sin embargo, no se tradujo en que apertura fuera el mejor predictor del desempeño académico en este meta-análisis. El hecho de controlar las relaciones de los Cinco Grandes con el desempeño académico por la relación del desempeño académico con la inteligencia, tuvo cierto efecto en la tabla de relaciones. En concreto, la puntuación en apertura disminuyó de .12 (para todos los niveles) a .09, aunque la relación con conciencia aumentó, debido a su negativa relación con inteligencia (-.03).

Poropat (2009) analizó también la validez añadida de los Cinco Grandes y de inteligencia sobre el desempeño académico en la escuela secundaria. En la literatura se encontró una relación relativamente alta ente la nota media de secundaria y académica que varía de .35 a .46 dependiendo del estudio (Kobrin et al., 2008; Kuncel et al., 2007; Robbins, Allen, Casillas, Peterson y Le, 2006; Schuler, Funke y Baronboldt, 1990). En el presente meta-análisis esta fue de .35. Controlando por el efecto de la nota media de secundaria, la relación de los Cinco Grandes con el desempeño académico en el nivel universitario el resultado fue .05 para amigabilidad, .17 para conciencia, -.01 para estabilidad emocional, .00 para extraversión y .03 para apertura. Respecto a la relación de inteligencia con el desempeño académico controlado por la relación del mismo con el desempeño en el nivel de secundaria, esta relación fue de .14. La conclusión que proviene de este análisis es que con la excepción de conciencia, los Cinco Grandes factores tienen poca o ninguna validez añadida al desempeño académico universitario cuando esta relación se controla por la nota media de la escuela secundaria. Del mismo modo, puede concluirse que conciencia tiene validez añadida en cuanto al desempeño académico ligeramente superior que inteligencia cuando se controla por la nota media de la escuela secundaria, lo que otra vez confirma su importancia como un predictor del desempeño académico.

En la Tabla 1 se encuentra un resumen de todos los meta-análisis respecto a la relación de la personalidad y el éxito académico que utilizaron como marco el modelo de los Cinco Factores. Como puede observarse, en cuanto al desempeño en el nivel de la enseñanza superior, hay ligeras diferencias respecto a los factores de amigabilidad, estabilidad emocional, apertura y extraversión, pero la relación del desempeño académico con conciencia siempre se mantiene en el nivel moderado, siendo el coeficiente promedio (.25). El máximo coeficiente encontrado fue .28 (Poropat, 2009, corregido por la restricción del rango; Salgado, 2000) y el mínimo el .22 (Poropat, 2005). La diferencia entre ambos valores puede deberse a que el primero es un coeficiente corregido, mientras que el segundo es la validez observada. Resalta por lo tanto que a pesar de las diferencias entre todos los meta-análisis (diferentes tamaños de muestra, diferentes correcciones aplicadas, etc.) sus resultados son muy parecidos (SD=.03). Además, en el estudio de Poropat (2009) se halló que la relación entre el factor de conciencia y el desempeño académico no está moderada por la edad del alumno, ni tampoco por sus capacidades cognitivas, y en el estudio de Trapmann y colaboradores (2007) que tampoco le influyen otras variables moderadoras tales como la carrera estudiada, la cultura, o la fuente de publicación. Finalmente, Poropat (2009) demostró también que conciencia tiene una validez añadida respecto al desempeño académico controlando por el desempeño en la escuela secundaria, siendo ésta superior a la validez añadida de inteligencia sobre el mismo también controlando por desempeño en la escuela secundaria.

Tabla 1. El resumen de los estudios meta-analíticos existentes sobre la relación entre los Cinco Grandes y el desempeño académico.

Estudio	Medidas del criterio	Nivel académico	Conciencia			Estabilidad Emocional[a]			Extraversión			Apertura a la Experiencia			Amigabilidad		
			k	N	r	k	N	r	k	N	r	k	N	r	k	N	r
Salgado (2000)[bcd]	calificac., calificac. medias	superior	15	2,214	**.28***	14	3,784	-.02	12	3,585	**.10***	11	1,654	**.00***	11	1,654	**-.04**
Poropat (2005)	calificac., calificac. medias	superior	16	4,626	**.22***	16	4,626	.03*	16	4,626	**-.04***	16	4,626	**.04***	16	4,626	**.08***
Trapmann et al. (2007)[bce]	calificac., calificac. medias	superior	41	10,855	**.27***	59	14,653	.00	54	12,424	**-.05***	41	14,942	**.13**	34	12,452	**.06**
O'Connor y Paunonen (2007)[b]	calificac., calificac. medias, medidas del desempeño en el aula y escritas	superior	23	5,879	**.24***	21	5,091	.03	22	5,161	**-.05**	23	5,878	**.06**	19	4,490	**.06***
Poropat (2009)[bcf]	calificac., calificac. medias	superior	92	32,887	.28	80	28,967	-.01	78	28,424	-.01	77	28,471	.09	75	27,944	.08
Poropat (2009)[bcf]	calificac., calificac. medias	secundaria	35	31,980	.26	24	25,495	.01	25	25,648	-.04	25	25,909	.16	24	25,488	.07
Poropat (2009)[bcf]	calificac., calificac. medias	primaria	8	3,196	.34	8	3,196	.24	8	3,196	.23	8	3,196	.32	8	3,196	.38
Poropat (2009)[bcf]	calificac., calificac. medias	mixto	135	68,063	.29	112	57,658	.02	111	58,518	.00	110	58,739	.16	107	56,628	.09
Trapmann et al. (2007)[bce]	abandono universitario	superior				4	2,714	**.06**	5	3,916	**-.01**	4	2,194	.03			
Trapmann et al. (2007)[bce]	satisfacción	superior				8	1,512	**.37***	4	660	**.10**						

Nota: k = número de estudios incluidos en la muestra, N = tamaño de la muestra, r = coeficiente de correlación, [a] los valores para el factor de neuroticismo han sido invertidos, [b] los coeficientes han sido corregidos por la falta de fiabilidad en el predictor, [c] los coeficientes han sido corregidos por la falta de fiabilidad en el criterio [d] los coeficientes han sido corregidos por la restricción del rango, [e] los coeficientes han sido corregidos por el error de muestreo, [f] los coeficientes han sido corregidos por la restricción del rango encontrada por Salgado (2000), * el intervalo de confianza no contiene el cero. Con negrita se han marcado los coeficientes para los que se conocen los intervalos de confianza.

Las facetas como predictores de las calificaciones académicas

En cuanto a la relación de las facetas con el desempeño académico, ningún meta-análisis existente ha incluido datos cuantitativos que permitiesen sacar conclusiones generalizables. Existen, sin embargo, algunos estudios individuales que trataron de explorar la naturaleza de esta relación, y comparar las capacidades predictivas de facetas frente las capacidades predictivas de los factores en cuanto a este criterio. A continuación, estos estudios se revisan en detalle, indicando en cada caso la metodología utilizada y las limitaciones de la misma, para poder responder a las cuestiones relacionadas con el dilema amplitud-fidelidad en el ámbito del desempeño académico.

En uno de los primeros estudios que abarcaban este problema, Rothstein y colaboradores (1994) examinaron si las facetas o los factores medidos por el PRF (D. N. Jackson, 1984) son mejores predictores del desempeño de los estudiantes en un curso de postgrado. Como criterio utilizó la nota media alcanzada al finalizar el curso, al igual que dos componentes específicos de la misma: desempeño en las pruebas escritas y el desempeño en el aula. Encontraron que en cuanto a la nota media, solo dos factores fueron predictores generalizables: amigabilidad (-.19) y apertura (.12), con los coeficientes de validez más bien pequeños a moderados. En cambio, una faceta de conciencia, el logro, y otra de apertura, dominancia, resultaron ser mejores predictores del desempeño que los factores a los que corresponden (.21 y .22 respectivamente). En cuanto a los restantes criterios del desempeño, ninguno de los factores predijo el desempeño en pruebas escritas, aunque varias facetas alcanzaron coeficientes de correlación significativos, pero muy pequeños (entre .09 y .10). Analizando la relación entre los factores y el desempeño en el aula, tres factores alcanzaron los coeficientes de validez predictiva significativos: extraversión .19, amigabilidad -.20 y apertura .17. No obstante, otra vez resultó que varias facetas superaron las capacidades predictivas de los factores del orden superior, como en el caso de amigabilidad su faceta de dominancia (-.23), en el caso de extraversión su faceta de exhibición (.33). Finalmente, una faceta de conciencia, el logro, resultó

predecir este tipo de desempeño (.21), aunque el factor del orden superior no era un predictor significativo.

Respecto a los resultados provenientes del estudio arriba mencionado cabe destacar que de ningún modo pueden considerarse definitivas. Es importante resaltar que el PRF (D. N. Jackson, 1984) no fue creado específicamente para reflejar el modelo de los Cinco Grandes, por lo que la correspondencia de los factores no puede ser exacta. Además, como ya se señaló en otro lugar, las escalas de PRF no pueden considerarse las facetas puras, ya que observando la estructura factorial de este inventario, vemos que presentan altas cargas factoriales en más de un factor. Todo esto indica que son más bien unas variables compuestas. Finalmente, para saber su validez específica (la que se debe únicamente a la faceta) habría que descontar de la varianza de las mismas esta parte de la varianza que se debe a los constructos de orden superior que las subyacen.

En un estudio posterior, De Fruyt y Mervielde (1996) encontraron, basándose en 30 facetas del NEO PI-R (Costa y McCrae, 1992b), que algunas predicen desempeño académico mejor que los factores a los que corresponden. Los criterios utilizados fueron la nota alcanzada en la primera convocatoria de los exámenes, la nota media alcanzada al final del curso, y el número de exámenes repetidos a lo largo del curso. Así, el valor de neuroticismo para predecir la nota media en la primera convocatoria de exámenes era -.16, pero las facetas de impulsividad y vulnerabilidad resultaron alcanzar el coeficiente de .22 en el caso de los hombres. La extraversión no resultó ser un predictor significativo del desempeño académico, aunque la faceta de actividad si lo era en el caso de todos los criterios utilizados alcanzando los coeficientes que varían de .15 (valor absoluto) al .26. Del mismo modo, aunque amigabilidad no mostró ninguna relación significativa con el desempeño, las facetas de confianza y franqueza eran predictores válidos del número de los exámenes repetidos en el grupo de los estudiantes masculinos. Igualmente, las facetas de apertura: fantasía y estética superaron a la validez predictiva del factor, la primera en dos casos, y la segunda en un caso. Finalmente, en cuanto al criterio de las calificaciones de la primera convocatoria, dos facetas de conciencia, autodisciplina y necesidad del logro, superaron a las propiedades predictivas del factor en el grupo de los hombres (.46 y .37 frente a .35 en el caso del factor), al igual que en el caso de las calificaciones alcanzadas al final del

curso (.36 y .30 frente a .28 en el caso del factor) y en el caso del número de exámenes repetidos -.36 y -.28 frente a -.25 en el caso de factor), siempre en los grupos masculinos. Pero, la conclusión de que en algunos casos las facetas predicen mejor el desempeño académico que los factores sería precoz, debido a que de la varianza de las facetas no se descontó en este estudio la comunalidad y la varianza debida al error.

El siguiente estudio que analizaba la cuestión de las propiedades predictivas de los factores y las facetas en cuanto al desempeño académico fue el de Paunonen (1998). Este estudio ya ha sido mencionado con anterioridad en el apartado dedicado a la revisión de estudios primarios relevantes para el dilema de fiabilidad-amplitud. En este punto se presentarán únicamente los resultados detallados encontrados para el criterio del desempeño académico definido como nota media académica. En la primera de las dos muestras Paunonen encontró que de los grandes factores únicamente amigabilidad fue un predictor del desempeño académico generalizable (-.24), siendo los factores medidos por el NEO-FFI (Costa y McCrae, 1992b). A su vez, resultó que un mayor número de facetas resultaron ser predictores generalizables, siendo estas medidas por el PRF (D. N. Jackson, 1984). Así, las facetas de inferioridad (*abasement*) y de defensa (*defendance*) correspondientes al factor de amigabilidad predijeron la nota media con los coeficientes de validez de -.21 y .21 respectivamente, logro como una faceta de conciencia .27, y finalmente, las facetas de juego (*play*) y de necesidad de ser apoyado (*nurturance*) correspondientes al factor de extraversión, coeficientes de -.27 y -.25 respectivamente. En la segunda muestra, ninguno de los factores medidos por el NEO-FFI resultó ser un predictor generalizable de la nota media académica, y de las 16 facetas, medidas por el JPI (D. N. Jackson, 1994), solo la responsabilidad resultó ser un predictor generalizable de este criterio .23.

Aunque Paunonen (1998) concluye que sus resultados demuestran que los rasgos específicos tienen capacidades predictivas superiores a los rasgos del orden superior, estas conclusiones no son del todo correctas. Primero, cabe recordar que se utilizaron medidas más largas para la medición de las facetas que para los factores, lo que no es correcto teniendo en cuenta que las facetas se encuentran en el nivel inferior que los factores en la jerarquía de personalidad, por lo que las medidas de factores, al componerse de varias facetas, lógicamente deberían ser más amplias que las primeras (véase pág. 46 del presente trabajo). Otra limitación es que las facetas de

PRF (D. N. Jackson, 1984) y del JPI (D. N. Jackson, 1976, 1994) fueron consideradas las facetas puras, siendo, en realidad, variables compuestas. Por ejemplo, la faceta de inferioridad del PRF perteneciente a amigabilidad se relaciona con este factor .36, pero tiene una segunda carga factorial de .27 en el factor de extraversión. Igualmente la necesidad de ser apoyado se considera una faceta del factor de extraversión (.44), pero su segunda carga factorial es también muy alta (.37 en el factor de amigabilidad). El caso de la escala de responsabilidad de JPI es muy similar; relacionándola con los factores del NEO-FFI resulta que tiene una carga de .36 en el factor de amigabilidad y otra de .24 en el factor de conciencia. Finalmente, cabe destacar el hecho de que para investigar la validez añadida de los factores sobre las facetas y viceversa, se utilizó el método de regresión múltiple que en sí no descuenta la varianza común de las facetas y la varianza del error de la varianza total de las mismas, por lo que su uso no ha sido correcto para sacar las conclusiones pertinentes en cuanto superioridad de factores o de facetas como predictores.

En otro estudio, Paunonen y Ashton (2001) compararon la capacidad predictiva de dos factores, conciencia y apertura a la experiencia, con la de las dos escalas específicas (una por cada factor) que, según los jueces, más se relacionan con el desempeño académico. En el caso de conciencia la escala elegida fue logro y en el caso de apertura a la experiencia fue comprensión (*understanding*); como criterio se utilizó la nota media académica. Se encontró que conciencia predice las calificaciones con un coeficiente de validez de .21. La capacidad predictiva de la faceta logro fue ligeramente superior, alcanzando .26. Sin embargo, las diferencias entre ambos coeficientes no fueron significativas. En cuanto a apertura a la experiencia, ésta mostró una relación con el desempeño académico de -.04 (no significativo), mientras que la faceta comprensión una validez de .23. Estos resultados llevan a los autores a concluir que los rasgos primarios tienen una mayor capacidad predictiva que los factores. Cabe resaltar que la metodología de este estudio muestra las mismas limitaciones que las que se indicaron para el estudio anterior, por lo que los resultados de Paunonen y Ashton no pueden considerarse generalizables.

En otro estudio, Gray and Watson (2002) analizaron la relación entre la nota media académica por un lado, y la nota media de la escuela secundaria por otro lado, con los Cinco Grandes medidos por el NEO-FFI, y en el caso del factor de conciencia por

el NEO PI-R (Costa y McCrae, 1992b). Conciencia, apertura y amigabilidad resultaron ser predictores válidos del desempeño académico alcanzando los coeficientes de .36, .18 y .15 respectivamente, en cuanto al desempeño en la secundaria solo conciencia fue un predictor válido .22. El análisis detallado de las facetas de conciencia indicó que aunque su poder predictivo no era significativamente superior al del factor, las facetas muestran relaciones diferenciadas con el desempeño, siendo orden la que alcanzó el coeficiente menor. 15 y la necesidad del logro el coeficiente más alto .39. En el caso del desempeño académico y en el caso del desempeño en la escuela secundaria cuatro de las seis facetas resultaron ser predictores válidos: responsabilidad, necesidad de logro, autodisciplina y logro alcanzando respectivamente los coeficientes de .27, .22, .19, .15. Estos resultados, sin embargo, no indican la superioridad predictiva de las facetas debido a que no se descontó del total de la varianza que explican esta parte de la varianza que las facetas deben a los factores del orden superior subyacentes y al error de medida.

El siguiente estudio que analiza los factores y las facetas como predictores del desempeño académico, en este caso de la nota media final de los estudiantes de medicina al final de los tres primeros años de carrera, es el de Lievens, Coetsier, De Fruyt y De Maeseneer (2002). Para medir el modelo de los Cinco Grandes se utilizó el NEO PI-R (Costa y McCrae, 1992b). Los resultados indicaron que el único factor que predijo desempeño de forma generalizable y través de todos los años de carrera fue conciencia; extraversión resultó ser un predictor del desempeño en el primer curso. A continuación se examinó la validez predictiva de las facetas pertenecientes a estos dos factores. Efectivamente, resultó que los estudiantes más propensos al fracaso académico puntúan más bajo en las facetas de conciencia: responsabilidad y necesidad de logro, y más alto en las facetas de extraversión: gregarismo y búsqueda de emociones. Esto podría sugerir que las facetas son mejores predictores de factores. No obstante, en el caso de las facetas de conciencia, de 12 casos cuando éstas resultaron ser predictores generalizables del desempeño, solo en dos casos su valor superó el valor alcanzado por el factor. Si añadimos a esto que en este estudio no se utilizaron los métodos que permitiesen estimar cuanta varianza se debe únicamente a la faceta en sí, podemos concluir que otra vez el dilema amplitud-fidelidad no ha sido resuelto.

En un estudio posterior, Chamorro-Premuzic y Furnham (2003b) también hallaron que algunos rasgos primarios predicen el desempeño académico (la nota media) mejor que los factores. El compuesto de los cinco factores (medidos por el NEO PI-R, Costa y McCrae, 1992b) consiguió explicar el 15% de la varianza en el desempeño académico, siendo el factor más válido conciencia, seguido por el factor de extraversión. A su vez, las 30 facetas de este inventario consiguieron explicar un 28% de la varianza, siendo el cambio en R^2 significativo para las facetas sentido del deber (*dutifulness*), logro (ambas correspondientes a conciencia), y actividad (correspondiente a extraversión). Los autores resaltan, no obstante, que el aumento en la cantidad de varianza explicada por los rasgos específicos no es proporcionalmente significativo al número de los predictores utilizados. Se consiguió una media de 3% de la varianza explicada por cada factor y menos del 1% de la varianza explicada por cada rasgo específico. Sin embargo, al no descontar de las facetas la varianza que se debe a los factores del orden superior y al error, puede suponerse que estos resultados serían aun menos significativos si se aplicasen los métodos analíticos adecuados.

El estudio más reciente que analiza el poder predictivo de los factores y de las facetas en cuanto al desempeño académico, ha sido llevado a cabo por Lievens, Ones and Dilchert (2009). Se basa en una muestra de estudiantes de medicina y utiliza como criterio la nota media alcanzada al final de cada uno de los 7 años de carrera. El diseño longitudinal de la investigación permitió a los autores hacer comparaciones respecto al poder predictivo de las grandes dimensiones y de las facetas a lo largo de toda la carrera. Para medir personalidad se utilizó el inventario NEO PI-R (Costa y McCrae, 1992b). Los resultados de este estudio demuestran que la validez tanto de factores como de facetas varía significativamente dependiendo del curso académico, alcanzando el nivel más alto, cuanto mayor sea el curso de la carrera. Así ocurre en el caso de todos los factores, con la excepción del neuroticismo, cuya validez permaneció casi invariable y nula. Los autores observaron que mientras en el primer año de carrera solo algunos factores y facetas demostraron validez superior a .10, en el séptimo año fue ya la gran mayoría. Explican estas diferencias con el cambio en los requerimientos y tareas exigidas para aprobar cada año, estando los primeros cursos de medicina centrados más en la adquisición de los conocimientos teóricos, y los últimos más en la

demostración del conocimiento procedimental. El aumento promedio de la validez de facetas a lo largo de la carrera fue .18 para las facetas de extraversión, .15 para las facetas de apertura, .10 para las facetas de amigabilidad, y .21 para las facetas de conciencia.

Sin embargo, debido a que el objetivo de Lievens y colaboradores (2009) no consistía en estimar la validez añadida de las facetas sobre los factores ni al revés, no se sacan conclusiones en cuanto al dilema amplitud-fidelidad. De todas maneras, una mirada más próxima a los resultados encontrados en este estudio permite hacer algunas deducciones al respecto. Aunque es indiscutible que ciertas facetas superaron a los factores en cuanto a su poder predictivo, hay que tener en cuenta dos hechos importantes: (1) se analizaron seis facetas por factor durante siete años de carrera, lo que da la posibilidad a que en 42 ocasiones las facetas pudiesen haber superado significativamente la validez predictiva de los factores, sin embargo, en el caso de neuroticismo esto se ha dado solo en 7 casos, en el de extraversión en 11 casos, en el de apertura 9 veces, amigabilidad 11 veces, y conciencia 9 veces; (2) en varias ocasiones, una faceta diferente del factor en cuestión resultó ser el mejor predictor del desempeño académico, por ejemplo, en el caso del factor de amigabilidad, modestia fue el mejor predictor del desempeño en el segundo año académico, en el tercero lo fue actitud conciliadora, en el quinto franqueza y en el sexto, altruismo. Todo esto impide concluir que las facetas siempre son mejores predictores que los factores, porque aunque lo fueran en algunos casos, no se demostró un patrón estable de estas relaciones ni para un curso en particular, ni tampoco para toda la carrera. Además, entre todos los predictores, únicamente el factor de conciencia y cinco de sus facetas (con la excepción del orden) resultaron ser predictores significativos para todos los años de carrera médica, y ninguna de estas facetas había superado al factor como predictor en todos los años de carrera. También se encontró que la validez de conciencia iba en aumento desde .18 en el primer año hasta .45 en el séptimo, cosa que ocurrió también, aunque en menor medida, con las restantes dimensiones de personalidad, con la excepción del neuroticismo. Este hecho llevó a los autores a concluir que los rasgos de personalidad son mejores predictores del desempeño académico al final, que al principio de carrera. Sin embargo, debido a que este estudio se aplicó a una carrera tan particular como es medicina, habría que comprobar la

generalización de estos resultados en el caso de carreras diferentes. Por el momento, la investigación meta-analítica demostró que únicamente conciencia es un predictor del desempeño académico generalizable (O'Connor y Paunonen, 2007; Salgado, 2000; Trapmann et al., 2007) y estable en todos los niveles educativos (Poropat, 2009). Además, hasta el presente tampoco se han encontrado diferencias significativas entre el desempeño al principio y al final de carrera, siendo ambos predichos con los coeficientes de validez similares tanto por los predictores cognitivos (Kuncel y Hezlett, 2007) como no cognitivos (N. Schmitt et al., 2009; N. Schmitt et al., 2007). Para terminar, tampoco se puede olvidar el hecho de que en este estudio no se descontó la comunalidad y la varianza debida al error de la varianza de las facetas; de no ser así, se alcanzarían los coeficientes de validez de las facetas más pequeños. En conclusión, hasta que no se apliquen los análisis estadísticos oportunos, el dilema amplitud-fidelidad en el ámbito del desempeño académico sigue sin resolverse.

Los demás estudios encontrados que intentaban aclarar el dilema amplitud-fidelidad en cuanto al desempeño académico se basaban en las facetas que no estaban enmarcadas dentro del modelo de los Cinco Grandes. Por lo tanto, han sido excluidos de la revisión antes mencionada, debido a que fue imposible realizar un análisis de la amplitud del constructo al que nos estamos enfrentando. No podemos decir que un constructo sea una faceta únicamente porque no está incluido en el modelo de Cinco Grandes, o porque el instrumento de su medición contiene un número de ítems limitado. Por lo tanto, en la revisión anterior solo se han incluido los estudios que utilizaron la taxonomía de los Cinco Grandes y sus facetas para resumir los resultados. Aun así, cabe resaltar que ninguno de los estudios antes mencionados está libre de limitaciones metodológicas. Estas limitaciones son exactamente las mismas que en el caso de los estudios dedicados a la relación de las facetas con el desempeño laboral. Las facetas de personalidad a menudo fueron definidas de manera distinta, además de variar su número por factor, dependiendo del inventario de personalidad utilizado. En ningún caso ha llevado a cabo un estudio sistemático de la estructura de los Cinco Grades en el nivel de facetas. Además, algunos de estos estudios se basan en los instrumentos que no han sido específicamente diseñados para medir el modelo de los Cinco Grandes. Finalmente, para sacar las conclusiones no se han llevado a cabo los análisis adecuados, sobreestimando la varianza explicada por las facetas, al no

descontar de su varianza la varianza debida a la comunalidad y al error. Consecuentemente, los resultados respecto a la relación que tienen los rasgos primarios con el desempeño académico no son concluyentes. Cabe esperar, teniendo en cuenta las deficiencias antes mencionadas, que la relación de las facetas con el desempeño académico sería inferior a la encontrada hasta ahora. Por otra parte, los estudios meta-analíticos existentes han demostrado que el único factor generalizable que predice el desempeño académico con una validez significativa al nivel práctico es conciencia.

Después de esta revisión de la literatura existente respecto a la relación de los factores y las facetas con el desempeño académico, se formulan las siguientes hipótesis:

H1. De entre todos los factores de personalidad solo el factor de Conciencia muestra una relación positiva con las calificaciones académicas.

H2. Existe una relación positiva de las facetas de Conciencia con las calificaciones académicas.

H3. Existe una relación del factor de Conciencia con las calificaciones académicas, superior a la relación de las facetas de Conciencia con las calificaciones académicas.

H4. Las facetas de Conciencia no añaden validez a la del factor para predecir las calificaciones académicas.

Personalidad y desviación académica – resumen de la investigación e hipótesis

Respecto a la desviación académica, ésta suscitó el interés de los psicólogos organizacionales después de que varios autores sugirieron que puede estar vinculada con la contraproductividad laboral (p. ej., Blankenship y Whitley, 2000; Lawson, 2004; Lucas y Friedrich, 2005). Lucas y Friedrich (2005) basándose en una muestra de estudiantes con previa experiencia laboral, demostraron que las conductas contraproductivas académicas se relacionan con un test de integridad (.44, corregido), y con dos diferentes índices de conductas contraproductivas organizacionales (-.35

y -.45, ambos coeficientes corregidos). Similares resultados encontraron Marcus, Lee y Ashton (2007) también en dos muestras: una canadiense y otra alemana. Hallaron una relación de .54 y de .32 entre las conductas contraproductivas académicas y organizacionales para la primera y segunda muestra respectivamente. Además, descubrieron que las conductas desviadas académicas se relacionan con los tests de integridad y que esta relación es muy similar a la relación de las medidas de contraproductividad organizacional con los tests de integridad. Los coeficientes de correlación encontrados entre el constructo de integridad y las conductas desviadas académicas varían de .40 a .67, dependiendo de la muestra y del tipo de test de integridad (si fue basado en personalidad o manifiesto [*overt*]). Este resultado fue muy similar a la relación entre la integridad y contraproductividad organizacional que varía de .44 a .62. Todo lo anterior nos lleva a la conclusión de que las conductas desviadas académicas, efectivamente, están relacionadas con el constructo de contraproductividad organizacional.

En cuanto a los vínculos entre la personalidad y las conductas contraproductivas en el ámbito educativo, el meta-análisis de Trapmann y colaboradores (2007) analizó, entre otras, la relación entre los Cinco Grandes y el criterio del abandono universitario, que sería un tipo de conducta contraproductiva académica, similar a la rotación en el trabajo. Sin embargo, debido a un bajo número de estudios que incluyesen los datos al respecto, solamente se pudo llevar a cabo el análisis con los factores de estabilidad emocional y extraversión que alcanzaron los coeficientes de correlación próximos a cero y no generalizables (véase Tabla 1).

Existen varios estudios primarios que relacionaron la contraproductividad académica o alguna de sus facetas con los Cinco Grandes. Por ejemplo, C. J. Jackson, Levine, Furnham y Burr (2002) buscaron un modelo de regresión que explicase la contraproductividad académica y, entre otras variables, incluyeron las dimensiones de personalidad. En una muestra de 109 estudiantes universitarios encontraron que el único factor de personalidad significativo dentro de este modelo fue neuroticismo. No obstante, cabe destacar que para medir personalidad se utilizó *Eysenck Personality Inventory* (*EPQ;* Eysenck y Eysenck, 1975) que abarca únicamente tres factores: psicoticismo, extraversión y neuroticismo, por lo que no refleja el modelo de los cinco grandes factores, aunque los dos últimos factores forman parte del mismo. Además, si

tenemos en cuenta la investigación sobre la contraproductividad organizacional y la extrapolamos a la contraproductividad académica, los coeficientes de correlación más altos corresponden a los factores de conciencia, amigabilidad y estabilidad emocional, por lo que de los tres, el único factor incluido en este estudio fue neuroticismo (baja estabilidad emocional).

Farsides y Woodfield (2003) encontraron en una muestra de 432 estudiantes que las ausencias a las clases estuvieron relacionadas de forma significativa con dos de los cinco grandes factores, conciencia (-.16) y amigabilidad (-.14), midiendo los factores de personalidad mediante el NEO-FFI (Costa y McCrae, 1989). En otro estudio que analizaba las ausencias de los estudiantes, Oswald, Schmit, Kim, Ramsay y Gillespie (2004) hallaron, basándose en una muestra de 644 estudiantes de primero de carrera, que los predictores válidos resultaron ser los factores de extraversión .10 y de conciencia -.27, utilizando para la medición de los factores el inventario IPIP (Goldberg, 1999).

En otro estudio más reciente, Marcus y colaboradores (2007) analizaron la relación entre los Cinco Grandes y las conductas desviadas académicas en dos muestras de estudiantes: canadienses (N=171) y alemanes (N=114). Para medir el modelo de los Cinco Grandes se utilizó *HEXACO-Personality Inventory* (HEXACO-PI; Lee y Ashton, 2004; Lee y Ashton, 2006) que mide 24 facetas subordinadas a seis factores, de los que cinco se aproximan o corresponden a los Cinco Grandes. Los resultados de este estudio indican que de los Cinco Grandes el mejor predictor de la desviación académica es el factor de conciencia (-.49 y .-48 para las muestras alemana y canadiense respectivamente), seguido por el factor de emocionalidad (similar al factor de estabilidad emocional) que alcanzó el coeficiente de -.29 en la muestra canadiense, y el factor de extraversión que alcanzó el coeficiente de .13 en la muestra alemana.

No conocemos ningún estudio que haya analizado las relaciones entre la desviación académica y la personalidad en el nivel de las facetas de los Cinco Grandes. Incluso teniendo en cuenta las sugerencias de algunos autores de que la primera es una manifestación de la contraproductividad organizacional (p. ej., Blankenship y Whitley, 2000; Lucas y Friedrich, 2005), y extrapolando a la desviación académica los resultados encontrados para las conductas contraproductivas organizacionales, no pueden sacarse ningunas conclusiones generales. Nos enfrentamos a la escasez de

estudios y a las limitaciones metodológicas en los estudios existentes. A su vez, la revisión de los estudios que analizaron la relación entre la personalidad y la desviación académica/contraproductividad organizacional en el nivel de las grandes dimensiones tampoco nos ofrece una visión clara. Cada uno de los factores resultó predecir alguna faceta de la contraproductividad en por lo menos un estudio. Si tenemos en cuenta únicamente la contraproductividad general, los predictores válidos y negativos de este criterio fueron: amigabilidad (C. M. Berry et al., 2007; Salgado, 2002), conciencia (C. M. Berry et al., 2007; Dalal, 2005; Dudley et al., 2006; Marcus et al., 2007; Salgado, 2002), y estabilidad emocional (C. M. Berry et al., 2007; C. J. Jackson et al., 2002). Extraversión resultó ser un predictor positivo (Marcus et al., 2007). Los meta-análisis al respecto son muy escasos, y en cuanto al criterio de desviación académica, en realidad no existe ninguno que abarcase diferentes conductas desviadas a la vez. Cabe destacar que los resultados más consistentes se encontraron en las relaciones entre el factor de conciencia y contraproductividad organizacional/desviación académica, que alcanzó los coeficientes de correlación de moderados a altos en todos los estudios en los que se utilizó como predictor. Por todo lo anterior, se considera que:

H5. Existe una relación negativa de los factores de Conciencia, Amigabilidad y Estabilidad Emocional con las conductas desviadas académicas.

H6. Existe una relación negativa de las facetas de Conciencia, Amigabilidad y Estabilidad Emocional con las conductas desviadas académicas.

H7. Existe una relación positiva del factor de Extraversión con las conductas desviadas académicas.

H8. Existe una relación positiva de las facetas de Extraversión con las conductas desviadas académicas.

H9. Existe una relación de los factores de Conciencia, Amigabilidad, Estabilidad Emocional y Extraversión con las conductas desviadas académicas, superior al valor absoluto de la relación de facetas correspondientes a estos factores con el mismo criterio.

H10. Las facetas de Conciencia, Amigabilidad, Estabilidad Emocional y Extraversión no añaden validez a la de los factores correspondientes para predecir las conductas desviadas académicas.

Personalidad y las medidas específicas del desempeño académico - hipótesis

Debido a que las medidas específicas del desempeño académico consisten en pruebas repetitivas y con un límite de tiempo, para terminarlas con éxito consideramos que deberían ser importantes las cualidades tales como: el esfuerzo constante a lo largo de la tarea, la escrupulosidad, la máxima concentración y el control en las situaciones de estrés, ya que el éxito depende de un único intento. Por lo tanto, se espera que las personas que puntúen alto en conciencia, al poseer las cualidades de constancia y minuciosidad, tengan mayores posibilidades de obtener mejores resultados en este tipo de pruebas. Al mismo tiempo, se supone que también las personas emocionalmente estables conseguirán mejores resultados en este tipo de pruebas, debido a que son más pacientes y que su desempeño no se verá afectado por la presión del tiempo. Acorde con estas expectativas, y teniendo en cuenta que, hasta el presente, ningún estudio ha demostrado la superioridad predictiva de las facetas sobre los factores utilizando las técnicas analíticas adecuadas para solucionar definitivamente el dilema amplitud-fidelidad, se formulan las siguientes hipótesis:

H11. De entre todos los factores de personalidad, solo los factores de Conciencia y de Estabilidad Emocional muestran una relación positiva con las medidas específicas del desempeño académico.

H12. Existe una relación positiva de las facetas de Conciencia y de las facetas de Estabilidad Emocional con las medidas específicas del desempeño académico.

H13. Existe una relación del factor de Conciencia y del factor de Estabilidad Emocional con las medidas específicas del desempeño académico, superior a la relación de las facetas de Conciencia y de las facetas de Estabilidad Emocional con el mismo criterio.

H14. Las facetas de Conciencia y de Estabilidad Emocional no añaden validez a la de los factores para predecir las medidas específicas del desempeño académico.

CHAPTER 3

RESEARCH

Method

Sample

The sample consisted of 410 Spanish university students, 291 females (71%) and 119 males (29%). Out of 410 participants 406 delivered the academic records needed to calculate their GPA and to assign them to a college major. Those 406 students represented 25 different college majors, which, for the purposes of the investigation, have been divided into 16 homogeneous groups. The number of participants for each group and corresponding percentages are presented in Table 2. The largest groups of students were taking degree courses in biology and veterinary science, psychology, and law (68, 60 and 55 students respectively), whereas the smallest number of participants were studying nursing (5 people) and business majors (6 people). The remaining four students were excluded from all analysis that required GPA and type of college major data which were not available for them.

Table 2. Academic characteristics of participants.

	College Majors groups	*f*	%	Valid %	Cumulative %
Valid	Philology	12	2.9	3.0	3.0
	Teaching, pedagogy, social education	31	7.6	7.6	10.6
	Labor relations	39	9.5	9.6	20.2
	Pharmacy	17	4.1	4.2	24.4
	Psychology	60	14.6	14.8	39.2
	Law	55	13.4	13.5	52.7
	Dentistry, medicine	15	3.7	3.7	56.4
	Biology, veterinary	68	16.6	16.7	73.2
	History, art history, philosophy	11	2.7	2.7	75.9
	Physics, engineering	16	3.9	3.9	79.8
	Political science	15	3.7	3.7	83.5
	Chemistry	27	6.6	6.7	90.1
	Nursing	5	1.2	1.2	91.4
	Business	6	1.5	1.5	92.9
	Journalism	15	3.7	3.7	96.6
	Mathematics	14	3.4	3.4	100.0
	Total	406	99.0	100.0	
Lost		4	1.0		
Total		410	100.0		

Personality measures

In the present study, personality was measured by means of three different personality inventories developed to represent FFM: the NEO PI-R (Costa & McCrae, 1992b), the Five Factors Personality Inventory (*Inventario de Personalidad de Cinco Factores, IP5F*) (Salgado, 1996, 1998b) and the Big Five Factor Inventory (*B5FI*) (Salgado, 2006)**.**

The NEO PI-R (Costa & McCrae, 1992b). This personality inventory contains 240 items that measure BF personality dimensions: *neuroticism (N), extroversion (E), openness to experience (O), agreeableness (A)* and *conscientiousness (C)* and six facets per factor. To N correspond facets: *anxiety, angry hostility, depression, self-consciousness, impulsiveness, vulnerability*; to E: *warmth, gregariousness, assertiveness, activity, excitement-seeking, positive emotions*; to O: *fantasy, aesthetics, feelings, actions, ideas, values*; to A: *trust, straightforwardness, altruism, compliance, modesty, tender-mindedness*; and to C: *competence, order, dutifulness, achievement striving, self-discipline,* and *deliberation.* Each factor is measured with 48 items while each facet is assessed by 8 items. Responses are based on a 5 point scale from *strongly agree* to *strongly disagree*, and the middle response is *neutral*. The score for each item goes from 0 to 4 points, and depending on the item, score 4 may be assigned to the *strongly agree* or *strongly disagree* response, the response *neutral* always being scored with 2 points. Raw scores are then converted to T scores, with a mean of 50 and a standard deviation of 10.

Alpha coefficients for factor scales found in the American employment sample varied from .86 for A to .92 for N and the mean was .89 (Costa, McCrae, & Dye, 1991). In the Spanish employment sample, factor reliabilities ranged from .85 for E to .91 for C and the mean was .88 (Costa & McCrae, 2008). In the case of facet scales in American sample, Cronbach's Alphas fluctuated from .56 for *tender-mindedness* to .81 for *depression* and the mean was .70. In Spanish sample reliabilities ranged from .55 for *order* to .81 for *actions*, and the mean was .67. Specifically for N facet scales, reliabilities varied from .68 to .81 in American sample, with the average of .75, and

from .57 to .78 in Spanish sample. For E in American they ranged from .63 to .73 and the mean was .71, for Spanish sample these values were from .57 to .73. For O, in American sample the lowest reliability was .58, and the highest was .80, with an average of .71, in Spanish sample reliabilities ranged from .51 to .81. In the case of A facets, in American sample reliabilities went from .56 up to .79 with a mean of .68 and in Spanish sample from .58 to .75. Finally, for C they ranged from .62 to .75 and the average was .68 in American sample and in Spanish sample they fluctuated from .55 to .75. Two-year retest reliabilities ranged from .83 to .91 for domains and from .64 to .86 for facets (McCrae, Yik, Trapnell, Bond, & Paulhus, 1998).

The convergent validity of the NEO PI-R when compared to other FFM-based inventories has also been demonstrated. For example, John and Soto (2007) found that the relationships between the NEO PI-R and the *Big Five Inventory* (BFI; John, Donahue, & Kentle, 1991) and between the NEO PI-R and *Trait Descriptive Adjectives* (TDA; Goldberg, 1992) were .87 and .83 respectively for E, .99 and .83 for A, .97 and .91 for C, .97 and .76 for N, and .90 and .78 for O (all of them corrected). The relationships found between the NEO PI-R and the Five Factor Inventory (FFI) (Costa, Yang, & McCrae, 1998) were .71 for E, .52 for A, .65 for C, .57 for N, and .32 for O. In the case of BFQ scales (Caprara, Barbaranelli, & Borgogni, 1993) these relationships were: .78 for N, .67 for E, .68 for O, .58 for A, and .73 for C. These are only some of the examples, as many more studies have been conducted demonstrating the construct validity of the NEO PI-R.

The NEO PI-R correspondence to FFM was also confirmed by means of exploratory analysis. Costa et al. (1991) factored the 240 NEO PI-R items. When five varimax-rotated principal components were examined, they matched the five anticipated factors. Correlations between the factor scores and the N, E, O, A and C domain scales were .91, .89, .95, .95, and .89, respectively. Also each facet scale had its highest loading on the intended factor. Although in five cases large secondary loadings appeared (angry hostility, activity, actions, altruism and competence), they were to be expected.

The IP/5F (Salgado, 1996, 1998b). This personality inventory is made up of 200 items assessing BF dimensions: N, E, O, A and C, 40 items per dimension. To each factor, between five and seven Homogeneous Item Clusters (HICs) are assigned. The following HICs correspond to N: *upset and hostility, tension, sadness, anxiety, carefree, insecurity, control/relaxation*; to E: *assertiveness, sense of group belonging, reserved, sense of humor, leadership, egocentrism*; to O: *imagination and intuition, creativity, adventure, range of interests, unconventional*; to A: *honesty and sincerity, modesty, cooperation/competitiveness, tolerance, trust/distrust*; and to C: *labor effort, order and organization, precision and attention to detail, prudence, high performance, and strictness with norms of behavior.* Responses are based on a 3-point scale (where 1 = *I agree*, 2 = *I am undecided*, and 3 = *I do not agree*). The initial raw scores are added together and then converted into standardized scores expressed as point T (M = 50; SD = 10).

Previous studies have demonstrated that reliabilities for each factor are satisfactory and vary between .74 for A to .90 for N (Salgado, 1998b). With the exception of A, all factors possess reliabilities above .80, and the mean reliability for all dimensions is .83. In the case of HICs, Cronbach's Alphas vary from .33 to .84, and 60% of them are superior to .60. For N HICs reliabilities range from .48 to .71 and the mean is .65. For E they vary from .48 to .71 with a mean of .64. For O HICs reliabilities go from .40 up to .60, and the average is .50. For A HICs reliabilities are between .35 and .70, and the mean is .48. Finally, for C HICs reliabilities fluctuate between .33 and .79, while the average is .64.

IP/5F test-retest reliability was also demonstrated after time periods of one and two years. Test-retest reliability for N was .91 and .86 for one and two year periods respectively, for E .90 and .83, for O .79 and .69, for A .65 and .58, and for C .72 and .65.

All factors measured by the IP/5F demonstrated significant correlations with the same factors as measured by the NEO Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992b) which is the abbreviated version of the NEO PI-R. So, for neuroticism the criterion validity was .70, for E .88, for O and A .55, and for C .58 (Salgado, 1998b). The construct validity of IP/5F scales was also proved by contrasting them with the

Hogan Personality Inventory (HPI; R. T. Hogan, 1986; R. T. Hogan & Hogan, 1992) scales. The Hogan Personality Inventory measures seven scales, and previous investigations have demonstrated that six out of seven are related to the Big Five factors: the HPI *adjustment* scale corresponds to Big Five neuroticism or emotional stability, HPI *ambition* and HPI *sociability* corresponds to Big Five extraversion, HPI *likeability* matches up with Big Five agreeableness, *prudence* with Big Five conscientiousness, *intellectance* with Big Five openness to experience, while the school success scale is not defined as a Big Five personality factor (R. T. Hogan & Hogan, 2002). Salgado, Moscoso and Lado (2003) found that the observed correlation between the IP/5F neuroticism scale and the HPI adjustment scale was -.66 (-.75 corrected). For the IP/5F extraversion and HPI ambition and sociability scales correlations were .60 and .62 respectively (.69 and .74 corrected). In the case of IP/5F openness to experience and HPI intellectance the correlation was .69 (.85 corrected), between IP/5F agreeableness and HPI likeability it was .37 (.51 corrected), and between IP/5F conscientiousness and HPI prudence it was .49 (.67 corrected).

IP/5F adjustment to the Big Five model was also confirmed by means of exploratory and confirmatory factor analysis (Salgado, 1996). In the first analysis, five factors corresponding to Big Five dimensions were found and effectively all the HICs had significant loading on conceptually corresponding factors. However, in three cases high loadings were observed also in other than the expected factor, which may be due to measurement error or/and to significant relations between some clusters. The confirmatory analysis was applied to test which one out of the three alternative structural models fits best to the data from the IP/5F. The tested models were: (a) the null model where none of the specific scales load on a common latent variable, (b) the five orthogonal factors model, and (c) the five oblique factors model. The null model did not show a good fit, whereas the five orthogonal factors model showed slightly better fit to the data than the oblique model as far as the parsimony index was concerned. The rest of the model fit indexes showed very similar values for the orthogonal and oblique models. These results support the construct validity of the IP/5F personality questionnaire.

The B5FI (Salgado, 2006). This personality inventory mirrors the Big Five model and was developed in Spanish following the 18 facet structure proposed by Saucier and Ostendorf (1999). The B5FI includes 146 items in total, however, 20 items measure social desirability and do not count towards total factor scores. Each factor contains between three and four facet scales measured with 7 items. Three facets correspond to N (scored as ES): *irritability, insecurity* and *emotionality*, E consists of four scales: *sociability, unrestraint, assertiveness* and *activity*, O is composed of three scales: *intellect, imagination-creativity* and *perceptiveness*; A comprises four scales: *warmth, gentleness, generosity* and *modesty*, and finally, four scales correspond to C: *decisiveness, orderliness, reliability*, and *industriousness*. Responses are given on a 3-point scale, where 1 = *I agree*, 2 = *I am undecided* and 3 = *I do not agree*. Next, raw scores are converted to T scores with the mean of 50 and a standard deviation of 10.

Analysis of the psychometric properties of the B5FI demonstrated good reliability, especially considering the conciseness of this inventory. Alpha coefficients for factor scales fluctuate from .75 for A to .88 for C, with a mean of .80. Also facet scales possess reliabilities equal or superior to .70 in all cases, with a mean of .73. To be precise, for N and O scales Cronbach's Alphas range from .72 to .73, for E they vary from .73 to .78, for A from .70 to .74, and for C they fluctuate from .71 to .76.

Criterion measures

Academic performance definition. Academic performance is understood as the different cognitions, or emotions, actions and results displayed by a student, that lead to obtaining the best grades possible in a specific course or a group of courses corresponding to a particular year of a college major. These actions, cognitions, emotions and results are related to final productivity in the college major, that is Grade Point Average (GPA), obtained after having successfully passed in all subjects (Salgado, 2008).

According to this definition, if we think about academic performance in terms of performance during exams, oral presentations, academic essays, and team work, it becomes evident that this concept is the final product of multiple behaviors and abilities. According to Kuncel, Hezlett and Ones (2004) as well as Kuncel, Crede and

Thomas (2007), getting good grades is due not only to studying efficiently, but also to ability to manage conflicts, to work in a team, to search for information efficiently, to communicate successfully, and to avoid disruptive behaviors.

Comparison of academic performance and job performance. Due to this multi-faceted character of academic performance, Kuncel et al. (2007; 2004) concluded that the nature of academic tasks is very similar to the nature of job responsibilities. This assumption is very similar to the one derived from qualitative reviews carried out in 80s which supported the premise that well-performing students also have more chance of achieving success in the professional field (Baird, 1985; Samson et al., 1984). Nevertheless, some authors disagreed with these conclusions (e.g., Sternberg & Wagner, 1993). From this more skeptical point of view, academic tasks are precisely defined and have only one correct answer, while job tasks are complex and ambiguous. However, Kuncel et al. maintain that this vision does not reflect correctly the complexity of behaviors which lead a student to pass exams or accomplish other obligatory tasks, such as writing academic essays, or preparing and giving oral presentations. From this perspective, academic duties can be seen as complex and vaguely defined, with no universal response, and they often oblige the student to search for additional information and to demonstrate creativity.

Apart from the nature of academic tasks, some other important differences between work and academic settings should be highlighted. Undoubtedly, the most marked of these differences is the financial relationship between performance evaluators and the individuals being evaluated in the workplace and in educational settings (Poropat, 2005). While at the workplace employees receive financial payments for their performance, in academic settings, students provide payments to the academic institution, which may be direct (by payment of fees) or indirect (e.g., through current or future taxation). Differences in the processes by which performance is defined and assessed, are also considerable. In the case of job performance, a range of constructs which appear to be relevant to most, if not all, jobs and occupations have been identified. Examples of these generally applicable dimensions of performance include, among others, task performance (Borman & Motowidlo, 1993), contextual performance (Borman & Motowidlo, 1993), later on

relabeled as citizenship performance (Borman et al., 2001; Coleman & Borman, 2000), counterproductive work behavior (J. Hogan & Hogan, 1989), and adaptive performance (Pulakos, Arad, Donovan, & Plamondon, 2000). In academic settings, it seems that there has been no equivalent analysis of the nature of performance. Grades for individual courses and overall composites such as GPA appear to be virtually unchallenged as measures of academic performance. Although recently some different outcomes such as citizenship performance, self-evaluated performance, absenteeism or intent to quit have also been used as approximations to the academic performance construct (Oswald et al., 2004; N. Schmitt et al., 2009; N. Schmitt et al., 2007), it has been used more for investigation purposes than in academic practice, where grades and GPA-based measures unquestionably predominate.

In spite of the abovementioned differences between job and academic performance, both constructs are denominated as “performance”. This draws attention to the common belief of both laypeople and social scientists that job and academic performance have something in common (Kuncel et al., 2004; Sarmiento, Beale, & Knowles, 2007). A theory of performance proposed by Campbell, McCloy, Oppler and Sager (1993) advocates that performance measures need to be (a) behaviorally based and (b) totally under the control of the individual. Furthermore, they argue that performance does not need to be directly observable, for example in the case of cognitive behaviors it is not. However, performance must be relevant to organizational goals and measured in terms of individual proficiency, and this is the case of academic grades and GPA criterion. According to Poropat (2005), academic performance consists of behaviors which can be evaluated in terms of individual proficiency at standardized tasks or behaviors, such as oral and written examinations, essays, projects, etc., where external factors are not determining. He also emphasized that to the extent that grades, GPA and other formal assessment in academic settings result from a student’s endeavors and qualities rather than the evaluator’s partiality, organizational barriers or external influences, they largely fulfill the conditions for performance measures.

Use of academic performance in organizations. Keeping in mind the premise that academic performance is somehow similar to job performance, it is not surprising that GPA is commonly included as an item in biodata measures (Brown & Campion, 1994). Albrecht, Carpenter and Sivo (1994) surveyed 491 employers recruiting at American universities and 80.5% of them admitted using the minimum GPA as the main basis for future employment. Also Rynes, Orlitzky and Bretz (1997) found that GPA is relatively frequently used by enterprises as a criterion in the first selection stage, although its use is more common in the case of recently graduated candidates than in the case of experienced workers. Roth and Bobko (2000) emphasized that in the case of academic recruitment, GPA is actually used in all selection stages: as a screen for obtaining an interview, as a piece of information in the interviewing process, and when selecting college applicants. The study carried out by Cable and Judge (1997) demonstrated that GPA is somehow related to the interviewer's evaluation of the profile-fit of a candidate (.22), to the interviewer's hiring recommendations (.17), and finally to organizational hiring decisions (.11). Regarding the fact that transfer of learning from university to the workplace is actually limited (Haskell, 2001), all these facts may be surprising. However, a meta-analytical study of Roth et al. (1996) demonstrated that academic performance is in fact related to job performance, and this relationship is slightly above .30 (corrected). This coefficient is moderate but significant because it shows that GPA validity as a predictor of job performance is close to that of an unstructured job interview (.33) (McDaniel, Whetzel, Schmidt, & Maurer, 1994), the most common selection method.

Furthermore, Roth and Clarke (1998) found in their meta-analytical study that grades also predict starting salary (.20), current salary (.28), and salary growth (.07). A posterior meta-analysis of Strenze (2007) indicated in turn a somewhat lower relationship of .09 between High School GPA/class rank and income, however, high school grades/class rank were found to be a very good predictor of educational and occupational success (.53 and .37 respectively).

Finally, the results of the qualitative review of studies dedicated to academic performance by Pascarella and Terenzini (1991) are also worth mentioning. Pascarella and Terenzini found that GPA is related to job satisfaction, and this relationship varies

from .09 to .18, depending on the study. They also observed that GPA is related positively and persistently across the studies to promotions (.17) and to occupational status at the beginning of a career (.14), although these coefficients are rather small. However, any generalizations based on this review should be very cautious, as the relationships found by Pascarella and Terenzini reflect simply the average correlations found in the literature, without analyzing the possible moderating variables such as for example social and economic status which may be crucial (Sirin, 2005).

The abovementioned results of quantitative and qualitative reviews on academic performance demonstrate that in spite of its rather moderate relationship to organizational behaviors, it is a relevant concept for work and organizational psychology and possesses practical implications for the business world.

Academic performance measures in the research. Although today, as previously mentioned, grades and GPA measures dominate academic decision-making, for the purposes of research some other academic performance measures have been used. Our review of literature on academic success led us to the following list of more or less common academic outcomes:

a) First-year College GPA (FYGPA) (e.g., Allen & Robbins, 2008; Busato et al., 2000; Kahn, Nauta, Gailbreath, Tipps, & Chartrand, 2002; Kuncel & Hezlett, 2007; Oswald et al., 2004; Robbins et al., 2006; N. Schmitt et al., 2007; Zajacova, Lynch, & Espenshade, 2005).
b) The end of academic year GPA (e.g., De Fruyt & Mervielde, 1996; Di Fabio & Busoni, 2007; Harackiewicz, Barron, Tauer, & Elliot, 2002).
c) Overall college GPA (GPA) (e.g., Chamorro-Premuzic & Furnham, 2003a; Goff & Ackerman, 1992; Harackiewicz et al., 2002; Kuncel & Hezlett, 2007; Musgrave-Marquart, Bromley, & Dalley, 1997; Noftle & Robins, 2007; Tross, Harper, Osher, & Kneidinger, 2000).
d) Final grades in a particular undergraduate course (e.g., M. O. Borg & Shapiro, 1996; Harackiewicz et al., 2002; Jones, 2008; John W. Lounsbury, Sundstrom, Loveland, & Gibson, 2003; Paunonen & Ashton, 2001; Robbins et al., 2006; Rothstein et al., 1994).

e) Graduate status/degree completion (whether student has graduated or not on the officially anticipated date) (e.g., Kuncel & Hezlett, 2007; N. Schmitt et al., 2009).
f) Number of re-exams (e.g., De Fruyt & Mervielde, 1996).
g) Research productivity (e.g., Kuncel & Hezlett, 2007).
h) Citation count (e.g., Kuncel & Hezlett, 2007).
i) Faculty ratings (e.g., Kuncel & Hezlett, 2007).
j) Retention (the length of time a student remains enrolled at an institution, until the completion of the program of study) (e.g., Robbins et al., 2004).
k) Counterproductive academic behaviors (e.g., K. J. Chapman, Davis, Toy, & Wright, 2004 [cheating]; Farsides & Woodfield, 2003 [absences]; Jones, 2008 [negative academic behaviors]; Oswald et al., 2004 [absenteeism]; N. Schmitt et al., 2009 [self-rated absenteeism]; N. Schmitt et al., 2007 [self-rated absenteeism and intent to quit]).
l) Professor or peer ratings of academic performance (Chamorro-Premuzic & Furnham, 2003a [academic behavioral indicators evaluated by tutors]; Farsides & Woodfield, 2003 [tutorial report grades]; Oswald et al., 2004 [peer-rated behaviorally anchored rating scales for multiple dimensions of college performance]; Rothstein et al., 1994 [classroom performance rated by instructors]).
m) Self-ratings of performance (normally on behaviorally anchored rating scales [BARS]) (e.g., Oswald et al., 2004; N. Schmitt et al., 2009; N. Schmitt et al., 2007).
n) Contextual performance, extra-academic activities (e.g., Farsides & Woodfield, 2003 [submission of non-assessed work]; Poropat, 2002 [citizenship performance peer-rated scale]; N. Schmitt et al., 2009; N. Schmitt et al., 2007 [self-rated citizenship behavior test]).

Scholastic, cognitive and non-cognitive predictors of academic performance. For predicting the foregoing academic outcomes scholastic, cognitive and non-cognitive predictors can be found in the literature. Among scholastic predictors the most common are:

a) High School GPA (HSGPA) (e.g., Gray & Watson, 2002; Noftle & Robins, 2007; N. Schmitt et al., 2009; N. Schmitt et al., 2007; Tross et al., 2000; Wolfe & Johnson, 1995).
b) High School rank (HSR) (e.g., Baron & Norman, 1992; Harackiewicz et al., 2002; Singleton, 2007; R. M. Smith & Schumacher, 2005).

In cognitive predictors we include:

a) Standardized scholastic aptitude tests, among others: *Scholastic Aptitude Test – SAT* (e.g., Baron & Norman, 1992; M. O. Borg & Shapiro, 1996; Harackiewicz et al., 2002; Noftle & Robins, 2007; Oswald et al., 2004; N. Schmitt et al., 2009; N. Schmitt et al., 2007; Singleton, 2007; R. M. Smith & Schumacher, 2005; Wolfe & Johnson, 1995), *American College Testing – ACT* (e.g., Bassiri & Schulz, 2003; Harackiewicz et al., 2002; Lindley & Borgen, 2002; Oswald et al., 2004; N. Schmitt et al., 2009; N. Schmitt et al., 2007), *Graduate Record Examination – GRE* (e.g., Anaya, 1999; W. R. Borg, 1963; Ji, 1998; Newman, 1968; Nilsson, 1995) and *Graduate Management Admission Test – GMAT* (e.g., Dobson, Krapljan-Barr, & Vielba, 1999; Graham, 1991; Nilsson, 1995).

Among non-cognitive predictors of academic performance one may find:

a) Personality dimensions (e.g., O'Connor & Paunonen, 2007; Poropat, 2009; Salgado, 2000; Trapmann et al., 2007).
b) Variety of psycho-social variables, among others: academic goals, institutional commitment, perceived social support, social involvement, etc. (e.g., Harackiewicz et al., 2002; Robbins et al., 2004).
c) Study habits, skills and attitudes (SHSAs) (e.g., Credé & Kuncel, 2008).
d) Situational judgment tests (SJTs) (e.g., Lievens, Buyse, & Sackett, 2005; Oswald et al., 2004; H. Peeters & Lievens, 2005; N. Schmitt et al., 2009; N. Schmitt et al., 2007).
e) Biodata (e.g., Oswald et al., 2004; N. Schmitt et al., 2009; N. Schmitt et al., 2007).
f) Assessment centers (e.g., Annen & Eggimann, 2006; Richter, 2003; Rode et al., 2005).

Reliability of various academic outcomes. Although there exist many studies which focus on academic performance, psychometric properties of academic performance outcomes have hardly been studied. If GPA is taken into account, Etaugh, Etaugh and Hurd (1972) found that FYGPA reliability varies from .69 to .81. Later on, Reilly and Warech (1993) found that overall college GPA reliability is .83, the value that has been also used in the meta-analysis by Roth, BeVier, Switzer and Schippmann (1996) and of Crede and Kuncel (2008) (in the last one also for first-semester freshman GPA and for FYGPA). At the same time, in the meta-analytical study by Robbins et al. (2004) and in Strenze (2007) it was assumed that GPA reliability is equal to 1.0 when this data is derived from university registers. However, this does not always seem to be true. Oswald et al. (2004) observed that grades from various courses do not always measure the same fundamental dimension of performance. For example, assessment tasks which contribute to grades and GPA can be substantially different in style (individual and group assignments, essays, one-choice or multiple-choice tests, oral exams, etc.) and may actually have little practical relation to each other, and even correlate differently with other variables. Furthermore, it should be highlighted that some studies used a GPA reported by the students themselves. Nevertheless, the correlation between official and self-indicated GPA is still high: a meta-analytical study by Kuncel, Credé and Thomas (2005) found a relationship between self-reported college GPA and GPA from academic transcripts of .90 (N=12,089, k=12). Although this relationship indicates high reliability, it is worth adding that only 54.3% of self-reported college GPA was accurate, while 34.5% was overestimated and 8.8% underestimated.

As far as the reliability of academic outcomes other than GPA is concerned, there is little data available. In the meta-analysis by Robbins et al. (2004), reliability of retention measures was assumed to be 1.0 because official university records were used. In some primary studies, individual Alpha coefficients can be found for other predictors. For example, Farsides and Woodfield (2003) observed that the reliabilities of professors' ratings of academic absenteeism and of academic performance were .85 and .83 respectively. At the same time, Chamorro-Premuzic and Furnham (2003a) reported on the reliabilities of various scales used for seminar leader's evaluation of

academic performance and found that they fluctuated between .59 (for "written expression") and .79 (for "participation"), with a mean of .68 for the eight scales used in the study. The reliability of self-ratings of performance (based on BARS) used in the studies by Schmitt et al. (2007) and by Schmitt et al. (2009) was .74 and .77 respectively, and in Oswald et al.'s (2004) study it was .80. At the same time, the reliability of the scale measuring students' intent to quit was .79 (N. Schmitt et al., 2007), and that of the citizenship behavior scale .85 (N. Schmitt et al., 2007) and .80 (N. Schmitt et al., 2009).

Relationships among different academic outcomes. Taking into consideration the fact that all the aforementioned academic outcomes are considered to be different "samples" of the same academic performance construct, one might expect these outcomes to be highly correlated. Nevertheless, the existing research demonstrates that these relationships are quite moderate, barely achieving coefficients above .50. Farsides and Woodfield (2003) found that tutorial report grades were related strongly to overall GPA (.46) and to absenteeism (-.59) as well as to voluntary work submission (the mean was .29). Overall GPA was related to non-assessed work submission with a coefficient of .15 on average. Oswald et al. (2004) reported correlation coefficients between FYGPA and academic absenteeism measures of -.53, between FYGPA and self-rated performance of .22, and between FYGPA and peer-rated performance of .29. They also found relationships of -.22 between self-rated performance and absenteeism and of -.16 between peer-rated performance and absences. Schmitt et al. (2007) also found that self-rated performance was negatively related to class absenteeism (-.27), and likewise to intent to quit (-.27), and positively related to citizenship behavior (.39) and FYGPA (.15). Citizenship behavior was at the same time negatively related to intent to quit (-.25). In the study of Schmitt et al. (2009) similar values were found for self-rated performance and absenteeism (-.30), and citizenship behavior (.50). Furthermore, a relationship of .22 was found between self-rated performance and overall GPA, and of .16 between the same criterion and graduation status. Finally, Jones (2008) found that negative academic behaviors such as turning work in late, skipping classes and procrastinating, were related to final course grade (-.46).

The above review on academic outcome and its predictors leads us to an important conclusion for our investigation. Moderate relationships between different academic outcomes suggest that studying GPA as the unique academic criterion may not be enough, as the academic performance construct seems to cover a wide variety of quite distinct behaviors. Therefore, not only GPA, but also alternative measures of academic performance should be used both in research and practice.

Grade Point Average is the broadest criterion in our study. Many authors define GPA as the most common academic performance measure (e.g., Kuncel et al., 2005; Poropat, 2009; Roth & Bobko, 2000). Its use is so common in scientific literature that very often it is simply used as a synonym of academic performance or academic success without further explanations (e.g., Bacon & Bean, 2006; Gray & Watson, 2002; O'Connor & Paunonen, 2007; Trapmann et al., 2007).

However, as previously mentioned, studying only GPA may not be enough to cover the multifaceted nature of academic performance. For example, Schmidtt et al. (2009) recommended "… to reconsider traditional GPA and graduation criteria and to expand the criteria space of college student performance" (pág. 1481) using alternatives to grades and to GPA academic performance measures. For that reason, as the second criterion we chose academic deviance. In the case of job performance, counterproductive behaviors were shown to be important, and to have both practical and theoretical value for organizations (J. Hogan & Hogan, 1989). We found it interesting to look at its role within academic settings, especially considering the fact that there have been no studies where this criterion was used as the academic performance dimension and, at the same time, was related to personality characteristics. Furthermore, the fact that academic deviance is a narrower academic outcome than GPA, will allow us further on to make bandwidth comparisons of academic performance criterion.

Finally, two narrow measures of academic performance combined into one score were used as the third criterion. Narrow measures evaluate delimited aspects of academic attainment, in this particular case, the ability for successful performance in administrative tasks which require concentration and attention. This criterion is the narrowest of all the criteria used in our study.

Measures based on GPA. The study participants were at various stages of completion of their degrees; however, all of them had finished their first year successfully. In order to obtain a comparable grade point average (GPA) for all the students, only first year grades (FYGPA) were taken into account, although we will henceforth refer to this criterion as GPA. GPA was calculated using students' academic transcript. GPA for all college majors was calculated according to the formula:

$$GPA = \frac{x_1 + x_2 + x_3 \cdots}{N}$$

where x_1, x_2, x_3… are final grades from single undergraduate first-year courses, and N is the total number of times a student has taken each first-year course exam (total number of failed exams plus total number of passed exams). Single grades were coded on a 4-point scale, where 0 = *no show* (*no presentado*) and *failure* (*suspenso*), 1 = *passed* (*aprobado*), 2 = *good* (*notable*), 3 = *very good* (*sobresaliente*), 4 = *outstanding* (*matrícula de honor*).

In order to obtain the reliability of the GPA used in our study, an internal consistency was calculated. As suggested by Bacon and Bean (2006) this is the most appropriate method for this type of research, because it explicitly analyzes the extent to which items within scales are related one to another. It is assumed that the grades obtained by a particular student all cover the same latent construct and therefore can be combined into a scale (in this case GPA). The internal consistency coefficient verifies whether grades are in fact related one to another, and whether they measure a broad construct of academic performance. Reliability is also a necessary, although not sufficient condition for further estimations of validity (Nunally, 1978).

In reliability calculations only grades from core degree subjects (not optional ones) were taken into account. For that reason, the analysis was carried out for every college major separately, as each of them follows a different academic program. Also additional grade coding was carried out to avoid the same score being assigned to students who did not show up for the exam and to those who failed. In this second coding (conducted only for reliability estimation purposes) value "0" was assigned to *no show*, "1" was assigned to *failure*, "2" to *passed* (*aprobado*) grade, "3" to *good* (*notable*), "4" to *very good* (*sobresaliente*) and "5" to *outstanding* (*matricula de honor*)

grades. After this second coding each grade was divided by the number of re-exams. For example, if the person obtained a *pass* grade, but attempted this exam twice, the corresponding value of three (two plus one) was divided by two, and the resulting value of one and a half was assigned to this particular first-year course. At the same time, a person that obtained *pass* grade in her/his first attempt would be assigned a value of two, and a person who did not show up for the exam a value of zero, etc.

For reliability analysis purposes we considered only college majors that were represented by at least 14 students, because in the others the number of students was too small to carry out this examination. The results of this analysis are shown in Table 3. The internal consistency coefficients found for grades corresponding to different college majors varied from .77 to .93, with an average of .87 (weighted). This value indicates good reliability and is slightly superior to the value used in the meta-analytical studies of Roth et al. (1996) and of Credé and Kuncel (2008).

Table 3. Mean reliabilities of GPA for college majors with N≥14.

College Majors	N	r_{xx}	SD
Pharmacy	16	.82	
Journalism	15	.77	
Political science	15	.88	
Mathematics	14	.85	
Physics	14	.93	
Chemistry	27	.93	
Labor relations	39	.85	
Law	53	.85	
Biology	67	.87	
Psychology	59	.91	
Weighted-sample average for natural sciences	138	.88	.05
Weighted-sample average for social sciences	181	.87	.05
Total weighted-sample average	319	.87	.05

Note: r_{xx} – reliability, SD – standard deviation. Natural sciences group includes Pharmacy, Mathematics, Physics, Chemistry and Biology Majors. Social sciences group includes Journalism, Political Science, Labor Relations, Law, and Psychology Majors.

As the next step, in order to compare the GPAs of students proceeding from different college majors, the GPA variable was standardized (so that it has a mean of 0 and a standard deviation of 1). This was done in two ways: first, according to the mean and standard deviation of a GPA within each college major group (GPA-I), and secondly, according to the general (university) mean and standard deviation (GPA-U). Descriptive statistics for both standardized GPA indicators and for both genders are presented in Table 4.

Table 4. Descriptive statistics for standardized GPA for male and female students.

	Gender	N	M	SD	SE
GPA-I	Women	290	-.057	.974	.057
	Men	116	.003	1.048	.097
GPA-U	Women	290	.004	1.031	.060
	Men	116	-.009	1.114	.103

Afterwards, the t-test was used to discover whether a significant difference existed between both standardized GPA indicators. Given that Levene's test has in both cases a probability greater than .05, we may assume that the population variances are relatively similar. Furthermore, the results of the t-test indicate that there are no significant differences between male and female students in the standardized GPA indicators (in the case of GPA-I: t (404) = -.545, p>.05, and for GPA-U: t (404) = .112, p>.05).

Measures based on academic deviance. Academically deviant behaviors were measured by means of a scale created specifically for academic settings by Salgado (2003a). It is composed of 30 items, with the responses going from 1 = *never* to 5 = *always*. The minimum score on this test is 30 and the maximum is 150. In general terms, this self-report questionnaire measures aspects of academic deviance such as: cheating, stealing and damaging other people's property, absenteeism, personal

norms and productivity/effort. Example items include *I consciously submitted poor quality works* and *I voluntarily skipped lectures and/or seminars.*

Narrow measures of academic performance. In this study, narrow measures of academic performance consisted of two time-limited, administrative assignments, namely: General Competency Tests GCT1 and GCT2. Scores from both narrow measures of performance were combined into one unique score by the addition of correctly solved items in both tests.

GCT1. This test consisted in comparing the original text with a copy, and finding mistakes in each test item (one line). Each line contained the name and last name of a person, his/her address and an amount of money associated with each person's bank account. The response was classified as correct when all mistakes were correctly found in a test item. The time was limited to three minutes. Afterwards, the total number of correctly solved test items was scored.

GCT2. In this test study participants were asked to assign to every person on the list his/her mail-box number, according to the first letters of the person's last name. For example, if last names starting with letters *Aa* to *Al* were included in the mail-box number 1, a person called *Abbot M. A*. should be assigned to this mail-box (number 1). The correct response consisted in writing the correct mail-box number at the side of the corresponding last name on the list. The time for this trial was limited to three minutes. The total number of correct responses was scored.

Procedure

We carried out this study during the years 2006 to 2008. We invited students who had successfully completed their first year at the University of Santiago de Compostela in Spain to participate. Responding to recruitment posters placed throughout University departments, the students were self-selected subjects.

Volunteers were asked to complete a three personality inventories, a counter productive behavior test, and two General Competency Tests. Afterwards they provided a copy of their academic record. Upon the completion of these tasks, the volunteers were paid 10 € for their participation.

Analytical methods

As we have widely discussed in the first part of this dissertation (see Chapter 2), the results of previous investigations on bandwidth-fidelity dilemma demand re-examination due to several inaccurate methodological decisions. The main limitation of previous investigations was the fact, that final conclusions, regardless of the method used to estimate the predictive validity of facets, were based on the observed correlations of personality factors or scales with criterions, instead of on residualized values for both factors and facets. In our study we corrected this error by using principal component analysis (PCA) to determine residualized weights for personality dimensions and for specific scales.

Another problem related to previous investigations is the fact, that when interpreting multiple regressions results, many authors drew their conclusions from the *squared sample multiple correlation*, known also as the *coefficient of determination* (R^2), or from the *squared population multiple correlation*, frequently called *the adjusted* R^2 ($\hat{R}^2$). However, neither coefficient is appropriate for resolving the bandwidth-fidelity dilemma, because they contain a considerable amount of bias. It is widely demonstrated that when one applies regression weights derived from one sample to a new one, they always yield lower than in the original sample R^2. The amount of this change is called *shrinkage* and is due to the fact, that regression

weights are optimized for values from the original sample. Therefore, the multiple regression equation will never fit in the same way to the new sample as it does to the one that it was derived from. Typically, the smaller the sample size, and the more predictor variables there are in the regression model, the greater the positive bias the R^2 contains. For that reason, correction formulas which penalize small sample size and a high number of predictors were developed. The result of these corrections is expressed as $\hat{R}^2$. It represents the multiple correlation produced in the population based on true regression weights; it means, the amount of variance that model would explain, if it was derived from the population rather than from the sample. However, this coefficient is not the proper one for conclusions regarding bandwidth-fidelity issues either, because we are interested in how well our model would perform as a predictor in a totally different set of data and not in the entire population.

As observed by Browne (1975) and later emphasized by Cattin (1980b), if research is theoretical, one may be interested in how well a set of predictors explains the criterion variable in the population. However, in applied research, it is more common to be interested in how well regression weights from the sample will serve as predictors for different samples from the same population. In other words, what we are interested in is the predictive validity of regression parameters, expressed by means of *the population cross-validated multiple correlation* ($\hat{R}_{cv}{}^2$).

Therefore, a cross-validated estimator has to be used if the purpose of the investigation is applied, as in our case. For this reason, the conclusions drawn from the multiple and hierarchical regressions in our study (the ones that use a high number of predictors) are based on $\hat{R}_{cv}{}^2$ value calculated by the means of the Browne-Cattin's formula (Cattin, 1980a, 1980b), instead of $\hat{R}^2$ or R^2 values, as in previous studies,. It is worth emphasizing, that the difference between the value of $\hat{R}_{cv}{}^2$ and $\hat{R}^2$ is almost null where simple regression is concerned; however, it becomes more significant when the number of predictors used in the regression increases.

Browne-Cattin's formula

Cattin (1980a, 1980b) modified Browne's equation (Browne, 1975) by dropping its second term. He argued that it always produces a very small value in comparison to the first term. For that reason, he suggested that the second term may be omitted without a substantial loss in precision. He also demonstrated, that his formula was less biased than others formulas developed with the same purpose. Cattin's modification of Browne's equation adopted the following form:

$$\hat{R}_{CV}{}^{2} = \frac{(N-k-3)\hat{R}^{4} + \hat{R}^{2}}{(N-2k-2)\hat{R}^{2} + k}$$

where *N* is the sample size used to estimate the equation, *k* is the number of predictors, and $\hat{R}^2$ for samples equal to or bigger than 50 is the squared population multiple correlation calculated by the means of Wherry's formula (Wherry, 1931). For samples smaller than 50 he provided a different formula; however, it does not apply to our study, as our sample is larger. Nevertheless, Cattin did not explain explicitly the meaning of parameter $\hat{R}^4$ included in his equation. As observed by Lautenschlager (1990, p. 85), many authors incorrectly interpreted this component as simply the square of $\hat{R}^2$, which leads to additional bias. He indicated, that the correct formula to be applied when calculating this component is the one presented by Browne (1975, p. 85) as Formula 3.2:

$$\hat{R}^{4} = \left(\hat{R}^{2}\right)^{2} - \frac{2k\left(1-\hat{R}^{2}\right)^{2}}{(N-1)(N-k+1)}$$

where $\hat{R}^2$ is obtained from Wherry's equation, ***N*** is the sample size and ***k*** is the number of predictors. According to Lautenschlager, the term being subtracted from the square of $\hat{R}^2$ substantially eliminates the bias present in the first component of the equation. We followed his indications when calculating $\hat{R}^4$ from Browne-Cattin's equation.

Results

Steps of analysis

Firstly, we conducted a PCA to verify if all the personality inventories used in our study yielded a five-factor structure. Secondly, we carried out a PCA for each of BF factors individually, using as variables the scales corresponding to the analyzed factor and proceeding from personality inventories employed in our study. Due to the fact, that each of these personality inventories contains a different number of item clusters (facets or HICs) per factor, for each of the factors we factorized from 15 to 17 scales. The aim of the PCA was to identify the common, as well as the unique variance of all the factorized scales. The first unrotated principal component (PC1), the one that explains the major part of the variability in the set of personality scales defining each of the BF factors, served as a measure of the corresponding BF factor. The rest of the unrotated PCs served as measures of facets. Third, we calculated the PC scores for all variables and for all study participants in each of the unrotated PC (based on their scores for the constituent variables). Applying this for all study participants we obtained their scores both in BF factors and in corresponding facets. Four, we carried out three types of regressions: (1) bivariate linear regressions to test the predictive properties of BF factors, (2) multiple regressions to test the predictive properties of BF facets, and (3) hierarchical regressions, where a BF factor was introduced in the first step and the corresponding facets in the second one, to verify whether facets show an incremental validity over factor when explaining the criteria of interest. Next, the *population cross-validated multiple correlation* was calculated by means of Browne-Cattin's equation (Cattin, 1980a, 1980b) both for multiple and for hierarchical regressions. As a last step, we compared regression results in order to test our research hypothesis (see Chapter 2).

Principal component analysis of personality inventories

The NEO PI-R (Costa & McCrae, 1992b). A principal component analysis was conducted on the 30 inventory scales with orthogonal rotation (varimax) and five factors were extracted. Six components had eigenvalues over Kaiser's criterion of unity (Kaiser, 1960), but following Cattell's (1966) scree test for the number of factors, it becomes apparent that only five first components should be retained, as the point of inflexion occurred at the sixth data point. Therefore, the NEO PI-R yielded a clear five-factor structure, and these five components explained 59% of the variance. Afterwards, factors were rotated to improve their interpretability. Table 5 shows components' loadings after rotation. Noticeably, out of the 30 NEO PI-R scales, 27 had the highest loading on the corresponding BF factor. Only three scales yielded higher loading on a component other than the corresponding BF factor. For example, N scale *impulsiveness* loaded higher on E and C (.48 and .36 respectively) than on the corresponding factor (.35). Also A scale *tender-mindedness* has a higher loading on O (.53) than on A (.45). Furthermore, E scale *assertiveness* had comparable loadings on three factors, on the corresponding E (.42), but also on N (.42) and on A (-.43). Similarly, O scale *feelings* loaded highly both on O (.54) and E (.54). Regardless of this, the mean scale loadings are still high and range from .59 for A to .75 for C, demonstrating the good construct validity of this inventory.

Table 5. NEO PI-R scale loadings after rotation.

	Components				
	PC1	PC2	PC3	PC4	PC5
NEO PI-R-ES1	-,137	**,837**	-,092	-,075	-,004
NEO PI-R-ES2	,006	**,629**	,018	,555	,056
NEO PI-R-ES3	,077	**,846**	,163	-,028	,001
NEO PI-R-ES4	,082	**,680**	,352	-,046	,152
NEO PI-R-ES5	**,359**	**,354**	**-,485**	,304	-,204
NEO PI-R-ES6	,232	**,817**	,043	-,071	,045
NEO PI-R-E1	,093	,154	**,717**	,359	,110
NEO PI-R-E2	-,103	,066	**,695**	,220	-,131
NEO PI-R-E3	,237	**,423**	**,423**	**-,436**	,070
NEO PI-R-E4	,190	,056	**,647**	-,230	,154
NEO PI-R-E5	-,290	,067	**,441**	-,292	,160
NEO PI-R-E6	-,063	,208	**,752**	,051	,189
NEO PI-R-O1	-,172	,059	,225	-,083	**,629**
NEO PI-R-O2	,085	-,070	,014	,004	**,763**
NEO PI-R-O3	,018	-,194	**,541**	-,100	**,545**
NEO PI-R-O4	-,116	,106	,208	-,064	**,542**
NEO PI-R-O5	,146	,189	-,106	-,037	**,731**
NEO PI-R-O6	-,052	,033	,044	,276	**,534**
NEO PI-R-A1	-,014	,278	,245	**,532**	,078
NEO PI-R-A2	-,027	-,153	,007	**,738**	,031
NEO PI-R-A3	,169	,124	**,395**	**,494**	,282
NEO PI-R-A4	,020	,055	-,134	**,742**	-,074
NEO PI-R-A5	-,038	-,314	-,048	**,583**	,003
NEO PI-R-A6	,021	-,110	,132	**,447**	**,532**
NEO PI-R-C1	**,770**	,256	,015	-,034	,102
NEO PI-R-C2	**,712**	-,152	-,038	-,071	-,130
NEO PI-R-C3	**,769**	,043	-,018	,259	,081
NEO PI-R-C4	**,787**	-,079	,169	-,240	,045
NEO PI-R-C5	**,826**	,176	,058	-,077	-,033
NEO PI-R-C6	**,663**	,120	-,367	,234	-,140

The IP/5F (Salgado, 1998b). Twenty nine scales of this personality inventory were factorized by means of PCA and five factors were extracted and rotated using the varimax method. This analysis revealed seven eigenvalues of the correlation matrix in excess of unity; however, based on the scree plot of eigenvalues a five-factor solution should be considered for factor retention, as the slope from factors 6 to 29 is virtually flat. The five-factor model explained 55.46% of the variance. The rotation showed that all but one scale had the highest loading on the corresponding personality factor (see Table 6). Only N facet *insecurity* loaded higher on E (.62) than on the corresponding factor (.50). However, E facets *leadership* and *egocentrism* also had secondary loadings on O (.47 and .37 respectively). In all cases, the average scale loadings were high and varied from .62 for A to .70 for C and O, thus confirming the construct validity of this inventory.

Table 6. IP/5F scale loadings after rotation.

	Components				
	PC1	PC2	PC3	PC4	PC5
IP5F-ES1	**,783**	,008	,012	,028	,136
IP5F-ES2	**,757**	,183	,123	-,001	-,196
IP5F-ES3	**,738**	,232	-,075	-,009	,018
IP5F-ES4	**,791**	-,005	,086	-,010	-,025
IP5F-ES5	**,526**	,117	-,081	-,152	,167
IP5F-ES6	**,495**	**,616**	,145	-,102	-,058
IP5F-ES7	**,722**	-,152	,030	,061	,148
IP5F-E1	,226	**,711**	,292	-,007	-,151
IP5F-E2	,125	**,709**	,222	-,062	,118
IP5F-E3	,008	**,807**	,113	-,182	,099
IP5F-E4	-,006	**,699**	,086	-,029	,130
IP5F-E5	,009	**,475**	**,473**	,139	-,314
IP5F-E6	-,068	**,404**	**,368**	-,063	-,233
IP5F-O1	-,122	,263	**,628**	,026	,017
IP5F-O2	,030	,192	**,762**	,149	-,063
IP5F-O3	-,069	,254	**,680**	-,151	,023
IP5F-O4	,078	,109	**,729**	,094	,072
IP5F-O5	,186	-,024	**,684**	-,113	-,049
IP5F-A1	,265	,204	-,040	,016	**,499**
IP5F-A2	-,057	-,096	-,226	,061	**,644**
IP5F-A3	,059	-,056	-,041	-,142	**,748**
IP5F-A4	,272	,038	,211	-,028	**,609**
IP5F-A5	-,157	,078	,056	-,250	**,599**
IP5F-C1	-,157	-,134	,198	**,635**	,085
IP5F-C2	-,066	-,013	-,128	**,757**	-,102
IP5F-C3	-,063	-,068	-,022	**,850**	,018
IP5F-C4	,221	-,202	-,219	**,631**	,078
IP5F-C5	,092	,125	,229	**,705**	-,197
IP5F-C6	-,049	-,043	-,010	**,630**	-,221

The B5FI (Salgado, 2006). Eighteen scales from the B5FI were factorized by the means of PCA to corroborate whether this inventory reflects the BF model. In this case, both Kaiser's criterion of eigenvalues superior to unity (1960) and Cattell's (1966) visual scree plot test for number of factors to be extracted overlap and suggest that five-factor structure is the most appropriate for extraction. Noticeably, these five factors explained 67.50% of variance, the highest proportion out of all personality inventories used in our study. Afterwards, the five factors were rotated using the varimax method and the resulting factors matrix is presented in Table 7. The vast majority of scales (15 out of 18) had the highest loading on the corresponding BF factor, and the means of these loadings range from .69 for C to .74 for A. However, a closer analysis of scales' loadings shows, that the scale *decisiveness* originally belonging to C loaded on this factor .36, whereas its loading on N is almost twice as high (.60). None of the previous factor analysis of the personality inventories showed a similar result for any of the scales. Normally, even though a scale does not have the highest loading on the corresponding factor, the difference is not so significant, rarely achieving more than .15. In this case the difference is .24. This would mean, that this scale does not actually belong to C factor, but rather is an N scale. For this reason, in posterior analysis we decided to include *decisiveness* in the N scales and to remove it from the C scales. After having eliminated this scale from the C scales, the average scale loading on this factor increased to .81, thus confirming the accuracy of our decision. Another two scales, that did not show the highest loading on the corresponding factors were not so problematic, as in both cases the difference between the primary and the secondary loading was minor. This was the case with N scale *irritability* whose primary loading was .56 on A (compared to .55 on the original factor), and of E scale *assertiveness,* which loaded slightly higher on N (.51) than on the corresponding E (.49).

Table 7. B5FI scale loadings after rotation.

	Components				
	PC1	PC2	PC3	PC4	PC5
B5FI-C1	,052	,320	**,596**	**,362**	,135
B5FI-C2	,014	-,005	,038	**,832**	-,093
B5FI-C3	,277	-,006	,122	**,754**	-,002
B5FI-C4	-,025	,055	-,039	**,830**	,207
B5FI-ES1	**,558**	-,145	**,552**	,016	-,074
B5FI-ES2	,172	,177	**,798**	-,081	,058
B5FI-ES3	-,112	-,164	**,821**	,026	-,019
B5FI-E1	,276	**,849**	,134	-,028	,130
B5FI-E2	-,253	**,828**	-,130	-,066	,025
B5FI-E3	,020	**,485**	**,514**	,295	,257
B5FI-E4	-,081	**,652**	,199	,169	,337
B5FI-A1	**,631**	**,486**	-,321	,035	,103
B5FI-A2	**,823**	,067	,019	-,011	-,021
B5FI-A3	**,737**	,082	,135	,180	,121
B5FI-A4	**,767**	-,273	,032	,079	-,082
B5FI-O1	-,059	,028	,006	,096	**,825**
B5FI-O2	,056	,193	,006	-,181	**,822**
B5FI-O3	,061	,192	,120	,280	**,451**

Separating factor and facet variance by the means of PCA

In order to estimate personality factor variance and personality facet variance from the set of personality scales, we used a PCA method. This analysis was conducted separately for each personality dimension and we factorized all scales corresponding to the analyzed factor from three personality inventories: the NEO PI-R (Costa & McCrae, 1992b), the IP/5F (Salgado, 1998b) and the B5FI (Salgado, 2006). The unique exception was B5FI scale *decisiveness*. According to Saucier and Ostendorf's (1999) 18 facet model, on which the B5PI was based, this scale pertains to the C factor; however, the previous PCA revealed that its loading on N/ES is almost twice as high as on C. Therefore, it was finally factorized among the N/ES scales.

As a result, for each factor we obtained from 14 to 17 unrotated principal components (depending on the total number of scales included in the analysis). The first unrotated principal component (PC1) resulting from each of these analyses, served as a measure of the corresponding personality factor, and the remaining unrotated principal components (from PC2 to PC_n) as the measures of the corresponding facets. Given below are the results of PCA for each of the personality factors.

Conscientiousness. In the case of C, we factorized 15 scales in total, 6 from the NEO PI-R, 6 from the IP/5F and 3 from the B5FI. As previously mentioned, the B5FI C factor originally contained four scales, but we omitted the first one, in line with the previous PCA of this questionnaire. The first component (PC1) resulting from a PCA of C scales, representing the C factor, explained 46% of the variance, and the remaining components (PC2 to PC15), corresponding to the unique variance of the C facets, explained the remaining 54% (see Table 8). Once again, it is important to highlight that this analysis gives preference to facets, as it assigns them more variance than they would normally have. The general rule in a principal components analysis is that all the components which are not significant and meaningful should be removed. However, we kept all the components in order to explain 100% of the variance. We can see in Table 9 that all C scales loaded significantly on the C factor (PC1). At the same time,

from the rest of the components representing the unique facet variance (PC2 to PC15), PC4 had only two significant loadings (superior to .40), each of PC2, PC3, PC5, PC6 and PC7 had a single significant loading, and the remaining eight components had none. Moreover, while the mean scale loading on the C factor (PC1) was of .67, in the case of significant loadings corresponding to the components from PC2 to PC15, only one was close to this mean, while in the rest of the cases they were much inferior.

Table 8. Eigenvalues, percentage of variance, and cumulative percentages for principal components of C scales.

Components	Initial Eigenvalues		
	Total	Variance %	Cumulative %
PC1	6,927	46,178	46,178
PC2	1,210	8,069	54,247
PC3	1,135	7,569	61,815
PC4	1,131	7,541	69,357
PC5	0,829	5,529	74,886
PC6	0,700	4,666	79,552
PC7	0,545	3,635	83,187
PC8	0,415	2,767	85,954
PC9	0,394	2,627	88,581
PC10	0,387	2,583	91,164
PC11	0,354	2,357	93,521
PC12	0,317	2,112	95,633
PC13	0,273	1,821	97,454
PC14	0,224	1,493	98,947
PC15	0,158	1,053	100,000

Table 9. Components matrix of C scales.

	Components														
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
B5FI-C2	**,701**	,118	-,426	,231	,111	-,077	-,112	,250	-,115	-,237	,157	-,132	,214	-,077	,049
B5FI-C3	**,521**	**,552**	-,393	-,042	,186	,232	,143	-,132	,343	,060	,050	,108	,050	,051	,005
B5FI-C4	**,685**	,292	-,246	-,167	,297	,038	-,302	-,007	-,293	,188	-,091	,006	-,212	-,007	-,043
IP5F-C1	**,526**	-,355	,164	-,111	**,600**	-,307	,193	,151	,137	,099	,077	,052	-,052	-,065	,022
IP5F-C2	**,746**	-,349	-,236	,349	-,106	,045	,025	-,126	,013	-,031	-,044	,016	-,187	,073	,265
IP5F-C3	**,766**	-,280	,079	,229	,021	,031	-,131	-,392	-,030	,144	,092	,011	,181	-,191	-,071
IP5F-C4	**,604**	,304	,470	,299	,058	-,146	-,220	-,007	,220	-,013	-,143	-,292	-,024	,068	,002
IP5F-C5	**,684**	-,186	,104	-,458	,049	-,034	-,238	-,149	,070	-,393	,058	,120	-,036	,115	-,051
IP5F-C6	**,544**	-,225	,319	-,001	,143	**,701**	,032	,186	-,023	-,022	-,038	-,037	,022	-,017	-,007
NEO PI-R-C1	**,718**	,166	,232	-,198	-,365	-,011	-,015	,102	,024	,150	,423	-,031	-,125	-,023	,031
NEO PI-R-C2	**,741**	-,295	-,272	,359	-,175	-,048	,109	,110	,046	,040	-,008	,007	-,086	,161	-,259
NEO PI-R-C3	**,716**	,235	,154	-,119	,045	-,046	,486	-,192	-,269	-,119	-,018	-,161	,005	,058	-,013
NEO PI-R-C4	**,732**	-,178	-,067	-,437	-,170	-,087	-,041	,077	-,011	,247	-,153	-,026	,265	,178	,075
NEO PI-R-C5	**,779**	,014	-,129	-,272	-,309	-,058	,095	,093	,127	-,075	-,257	-,010	-,088	-,299	-,023
NEO PI-R-C6	**,653**	,317	,379	,339	-,079	-,111	,012	,110	-,113	-,029	-,088	,395	,074	,008	,026

Neuroticism/Emotional Stability. For N (scored inversely as ES) we factorized 17 scales in total, 6 from the NEO PI-R, 7 from the IP/5F and 4 from the B5FI (three of them originally belonged to the N factor, and one to the C factor). The first component, resulting from a PCA and corresponding to the N/ES factor explained 41.4% of the variance (see Table 10), while the remaining components, representing the unique variance of the N/ES facets, explained the remaining 58.6%. It is worth pointing out, that this analysis gives a certain preference to facets, because all their variance is taken into account, whether it is significant and meaningful, or not. But even so, the results found for facets cannot be directly compared to the results found for N/ES factor. As may be seen in Table 11 which contains a components matrix, 16 out of 17 scales had loadings superior to .40 on the N/ES factor (PC1), and the mean of all loadings was .63. At the same time, from components representing the unique variance of the N/ES facets (PC2 to PC17), PC2 had five loadings superior to .40, each of PC3, PC4, PC5 and PC7 had two and PC9 had one. The rest of the unrotated principal components did not have any loading superior to .40. Moreover, all but one of the loadings corresponding to the components from PC2 to PC17 were inferior to .63, that is to the mean scale loading on the N/ES factor (PC1).

Table 10. Eigenvalues, percentage of variance, and cumulative percentages for principal components of N/ES scales.

Components	Initial Eigenvalues		
	Total	Variance %	Cumulative %
PC1	7,044	41,434	41,434
PC2	1,632	9,600	51,034
PC3	1,442	8,480	59,514
PC4	0,928	5,459	64,973
PC5	0,909	5,348	70,320
PC6	0,717	4,219	74,540
PC7	0,635	3,733	78,273
PC8	0,557	3,278	81,551
PC9	0,500	2,940	84,491
PC10	0,446	2,623	87,114
PC11	0,412	2,421	89,536
PC12	0,392	2,304	91,839
PC13	0,334	1,967	93,806
PC14	0,304	1,787	95,592
PC15	0,290	1,707	97,300
PC16	0,246	1,449	98,749
PC17	0,213	1,251	100,000

Table 11. Components matrix of N/ES scales.

	Components																
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16	PC17
B5FI-C1	,418	,030	**,617**	,234	-,063	,353	,422	-,017	,230	,052	-,055	-,070	,093	,062	,065	-,004	-,039
B5FI-ES1	,424	**,512**	,347	-,043	,467	-,253	,000	-,027	-,121	,099	,263	,031	,241	,007	-,002	,023	-,006
B5FI-ES2	**,597**	,183	,559	-,193	-,097	-,162	,038	,079	-,177	,119	-,092	-,093	-,334	-,202	-,030	,025	-,004
B5FI-ES3	,450	,427	,399	-,064	-,251	,206	**-,510**	,055	,103	-,199	-,042	,136	,035	,100	-,007	,018	,007
IP5F-ES1	**,727**	,142	-,260	,116	,105	-,081	-,184	-,111	,406	,030	-,051	-,223	,052	-,286	,036	,037	-,010
IP5F-ES2	**,748**	-,176	-,074	,176	-,217	-,016	,014	-,230	,058	,088	,330	,315	-,151	-,056	,166	,008	-,018
IP5F-ES3	**,758**	-,064	-,085	-,072	,067	,284	-,158	-,323	-,191	,269	-,096	-,089	,037	,055	-,044	-,233	,097
IP5F-ES4	**,740**	,059	-,120	-,026	-,344	-,314	,054	,092	,165	,088	,141	-,159	-,028	,247	-,206	-,111	-,088
IP5F-ES5	,488	,069	-,271	**-,649**	,078	,325	,155	,268	,128	,028	,138	,098	,008	-,082	-,038	-,031	,018
IP5F-ES6	**,635**	-,564	,128	-,012	,214	-,069	-,135	,151	,114	,176	-,043	,005	-,031	,169	,008	,225	,210
IP5F-ES7	**,627**	,375	-,266	-,159	-,242	-,213	,197	-,107	-,040	,102	-,345	,205	,144	,038	,076	,124	,009
NEO PI-R-ES1	**,785**	-,067	-,101	-,088	-,219	-,019	,054	,060	-,224	-,253	,110	-,289	,082	,032	,284	,016	,093
NEO PI-R-ES2	**,655**	,250	-,169	,025	,485	-,024	,128	-,132	,087	-,294	-,100	,039	-,265	,174	,024	-,059	,006
NEO PI-R-ES3	**,804**	-,297	-,098	-,017	,104	,196	-,073	-,084	-,191	-,037	-,009	-,057	,029	,016	-,100	,214	-,310
NEO PI-R-ES4	**,664**	-,478	,138	,073	,144	-,192	-,079	,291	-,002	-,037	-,164	,157	,103	-,062	,120	-,245	-,125
NEO PI-R-ES5	,342	,470	-,382	**,513**	,034	,204	-,026	,389	-,144	,174	,005	,002	-,078	,006	,024	,018	,010
NEO PI-R-ES6	**,799**	-,156	,006	,237	-,086	,003	,144	-,004	-,111	-,256	,018	,107	,100	-,165	-,313	,000	,169

Extraversion. For E, the total number of factorized scales was 16, 6 from the NEO PI-R, 6 from the IP5F and 4 from the B5FI. As is displayed in Table 12, the E factor (represented by PC1) explained 38.7 % of variance, while the unique variance of the E facets (represented by components from PC2 to PC16) was responsible for the remaining 61.3%. Even though the majority of variance was explained by the unique facet variance, we must observe that not all this variance is in fact significant. However, we preserved all the components obtained from a PCA in order to explain 100% of variance in terms of the E factor and its facets. As we can see in Table 13, our results show that all E scales loaded significantly on the E factor (PC1), and the average scale loading on this component was .61. At the same time, none of the remaining components (PC2 to PC16) representing the unique facet variance, showed a comparable result. To be precise, only four loadings superior to .40 corresponded to PC2, to PC3 three, and to PC5, PC6, PC7 and PC8 only one. No significant loading corresponded to the remaining nine components. In addition, none of these significant loadings on components PC2 to PC16 was superior to .61, that is to the mean E scale loading on the E factor (PC1).

Table 12. Eigenvalues, percentage of variance, and cumulative percentages for principal components of E scales.

Components	Initial Eigenvalues		
	Total	Variance %	Cumulative %
PC1	6,195	38,719	38,719
PC2	1,732	10,824	49,543
PC3	1,263	7,891	57,435
PC4	1,000	6,249	63,683
PC5	0,976	6,098	69,781
PC6	0,732	4,574	74,355
PC7	0,654	4,090	78,444
PC8	0,590	3,688	82,132
PC9	0,475	2,970	85,102
PC10	0,446	2,789	87,891
PC11	0,416	2,601	90,492
PC12	0,375	2,341	92,833
PC13	0,341	2,129	94,962
PC14	0,297	1,856	96,819
PC15	0,267	1,669	98,487
PC16	0,242	1,513	100,000

Table 13. Components matrix of E scales.

	Components															
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
B5FI-E1	**,694**	,242	,457	,049	,174	,117	-,171	,016	-,111	-,080	-,026	-,134	,095	-,159	-,100	-,306
B5FI-E2	**,580**	,336	,083	-,241	,458	,208	-,295	-,230	,037	-,079	,060	,096	,115	,121	-,031	,208
B5FI-E3	**,401**	**,476**	,470	-,136	-,100	-,090	,516	,092	-,009	-,052	,206	,160	,056	,079	-,001	,020
B5FI-E4	**,544**	**,572**	,287	,216	-,102	,033	-,084	,089	,009	,246	-,200	-,157	-,247	-,002	,127	,138
IP5F-E1	**,788**	-,012	-,170	-,088	-,210	-,026	,111	-,169	-,154	-,176	-,243	,044	-,208	,110	-,293	,004
IP5F-E2	**,742**	-,280	,009	,000	,235	-,337	,126	,008	,094	,116	,036	-,080	,044	-,315	-,156	,178
IP5F-E3	**,715**	-,372	,035	-,210	,154	,022	-,022	-,058	,200	-,042	,277	-,051	-,365	,046	,098	-,107
IP5F-E4	**,581**	-,278	,000	-,291	-,287	,470	,172	-,218	,086	,251	-,115	,069	,108	-,120	,069	-,023
IP5F-E5	**,616**	,267	-,461	,005	-,182	-,098	-,082	-,081	-,265	,285	,312	-,122	,077	,083	-,045	-,047
IP5F-E6	**,483**	,177	-,508	-,072	,350	,234	,110	**,495**	-,031	,031	-,078	,150	-,043	-,043	-,002	-,047
NEO PI-R-E1	**,645**	-,471	,200	-,103	-,106	-,003	-,043	,331	,118	,027	-,061	-,237	,183	,275	-,056	,044
NEO PI-R-E2	**,589**	-,452	,202	,227	,251	-,255	-,027	-,078	-,222	,189	-,131	,288	,021	,119	,129	-,069
NEO PI-R-E3	**,658**	,269	-,345	-,233	,019	-,324	,104	-,137	,091	-,194	-,192	-,132	,124	-,002	,255	-,072
NEO PI-R-E4	**,642**	,188	-,070	,179	-,416	-,148	-,321	,092	,342	-,020	,044	,284	,054	-,036	-,050	-,045
NEO PI-R-E5	**,421**	-,012	-,191	,745	,168	,219	,252	-,180	,183	-,073	,045	-,084	,065	,079	-,013	-,012
NEO PI-R-E6	**,708**	-,273	,037	,170	-,279	,160	-,092	,117	-,297	-,315	,119	,008	,011	-,132	,161	,154

Openness. For O, the total number of scales included in the PCA was 14, 6 from the NEO PI-R, 5 from the IP/5F and 3 from the B5FI. The first component, corresponding to the O factor level, explained 35.5% of the variance, while the remaining components (PC2 to PC14), representing the unique variance of the O facets, explained the remaining 64.5% (see Table 14). This analysis served to explain all variance in terms of the O factor and its facets; although, it does not necessarily mean that all obtained components are significant and interpretable. We basically aimed to obtain separate estimates of factor and facet variance to be able to compare later their predictive power. However, it should be highlighted, that this analysis favors to a certain extent facets, as it attributes more variance to them than they would normally have. Yet, as we may observe in Table 15, the results found for the O factor (PC1) are much more significant than those found for the O facets. First of all, each and every one of the O scales had substantial loadings on the O factor (PC1), with a mean of .58. Moreover, the highest loading of 11 out of the 14 scales included in the analysis corresponded to this component. In comparison, none of the rest of the unrotated principal components representing the unique facet variance had more than three significant loadings (superior to .40). To be exact, PC2 had three significant loadings, PC3 had two significant loadings, and PC5, PC6, PC7, and PC12 each had a single significant loading. Moreover, only in two cases were these loadings equal or superior to .58, that is the average scale loading on the O factor (PC1). The remaining seven unrotated principal components did not have any significant loadings.

Table 14. Eigenvalues, percentage of variance, and cumulative percentages for principal components of O scales.

Components	Initial Eigenvalues		
	Total	Variance %	Cumulative %
PC1	4,975	35,535	35,535
PC2	1,389	9,923	45,459
PC3	1,257	8,978	54,437
PC4	0,992	7,085	61,522
PC5	0,919	6,567	68,089
PC6	0,767	5,479	73,568
PC7	0,649	4,636	78,203
PC8	0,581	4,147	82,351
PC9	0,517	3,693	86,044
PC10	0,462	3,302	89,345
PC11	0,435	3,108	92,453
PC12	0,384	2,742	95,195
PC13	0,357	2,550	97,746
PC14	0,316	2,254	100,000

Table 15. Components matrix of O scales.

	Components													
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
B5FI-O1	**,495**	-,226	**,563**	-,326	,165	,238	-,182	,147	,128	,104	-,075	,265	,188	,064
B5FI-O2	**,635**	-,358	,239	-,044	-,014	,457	,108	-,024	-,059	-,045	-,068	-,414	-,078	-,030
B5FI-O3	,292	-,532	,060	,233	**,645**	-,241	,179	-,121	-,099	,050	,198	,029	-,006	,022
IP5F-O1	**,687**	-,336	-,137	,274	-,144	-,157	,159	,099	,185	-,022	-,325	,116	,011	-,296
IP5FI-O2	**,711**	-,218	-,325	-,096	-,235	-,083	,182	,043	-,149	,206	-,160	,073	-,087	,364
IP5FI-O3	**,573**	-,075	**-,582**	-,092	-,025	,230	-,150	,107	,169	,274	,332	,008	,000	-,133
IP5FI-O4	**,716**	,146	-,069	-,313	-,048	-,342	-,002	,057	-,239	-,035	,029	-,185	,374	-,084
IP5FI-O5	**,581**	,318	-,257	-,313	,276	,073	,297	,010	,288	-,365	,017	,042	-,034	,087
NEO PI-R-O1	**,632**	,045	,138	,371	-,367	,239	,150	-,177	-,179	-,200	,248	,217	,121	,008
NEO PI-R-O2	**,650**	,142	,273	,166	-,152	-,266	-,157	-,362	,397	,112	,047	-,143	,017	,115
NEO PI-R-O3	**,653**	-,019	-,080	,328	,096	-,089	-,491	,321	-,058	-,266	,000	-,046	-,094	,114
NEO PI-R-O4	**,553**	,421	-,219	,045	,336	,211	-,208	-,371	-,214	,092	-,248	,081	-,029	-,076
NEO PI-R-O5	**,665**	,146	,360	-,354	-,115	-,231	,009	,018	-,154	,020	,130	,082	-,377	-,160
NEO PI-R-O6	,302	**,647**	,262	,374	,168	,051	,268	,337	,007	,247	,005	-,062	,015	,017

Agreeableness. In the case of A, a PCA was carried out for 15 scales in total, 6 from the NEO PI-R, 5 from the IP5F and 4 from the B5FI. As shown by Table 16, PC1 representing the A factor underlying all scales, explained 31.5%, while components from PC2 to PC15 corresponding to the unique variance of the A facets explained the remaining 68.5%. Similarly, as in the case of other personality factors, this analysis was conducted to explain the whole variance in terms of the A factor and its facets; therefore, even insignificant and non-meaningful variance was taken into consideration. The components matrix (see Table 17) shows that all scales had loadings superior to .40 on the A factor (represented by PC1), and the mean loading was .56. At the same time, from the remaining components, PC3 had only three loadings superior to .40, PC2 had two and each of PC4, PC5, PC7 and PC8 had only one. No significant loadings corresponded to the remaining unrotated principal components.

Table 16. Eigenvalues, percentage of variance, and cumulative percentages for principal components of A scales.

Components	Initial Eigenvalues		
	Total	Variance %	Cumulative %
PC1	4,722	31,477	31,477
PC2	1,929	12,861	44,338
PC3	1,583	10,555	54,893
PC4	0,901	6,004	60,896
PC5	0,890	5,936	66,832
PC6	0,807	5,380	72,212
PC7	0,718	4,789	77,002
PC8	0,566	3,774	80,776
PC9	0,534	3,558	84,334
PC10	0,482	3,215	87,549
PC11	0,466	3,104	90,653
PC12	0,412	2,748	93,401
PC13	0,397	2,647	96,049
PC14	0,324	2,161	98,210
PC15	0,268	1,790	100,000

Table 17. Components matrix of A scales.

	Components														
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
B5FI-A1	**,416**	-,233	**,636**	,161	-,019	,260	-,179	,194	,300	,065	,224	,132	,064	-,193	,019
B5FI-A2	**,584**	-,367	,372	-,105	-,242	,253	-,044	,015	,064	-,281	-,177	-,256	-,084	,249	,026
B5FI-A3	**,539**	-,433	,317	-,301	-,045	-,331	-,029	-,011	-,125	,187	-,077	,248	,258	,180	-,042
B5FI-A4	**,610**	-,476	,010	-,121	-,258	-,189	,263	,063	-,156	,079	-,097	-,057	-,259	-,316	,018
IP5F-A1	**,415**	**,698**	,101	,035	-,197	-,271	,183	,193	,154	-,050	-,116	-,013	,128	-,004	,300
IP5F-A2	**,568**	-,027	-,485	,020	,083	,062	-,415	,402	-,099	,142	-,036	,079	-,196	,114	,089
IP5F-A3	**,679**	,009	-,286	-,071	,209	,244	,060	-,175	,270	,355	-,255	-,177	,128	-,050	-,018
IP5F-A4	**,475**	,370	,052	**-,554**	,212	,272	,216	,146	-,208	-,015	,288	-,101	,056	-,011	-,014
IP5F-A5	**,537**	,347	,173	,329	-,314	,298	-,028	-,301	-,344	,189	,062	,086	-,030	,051	,081
NEO PI-R-A1	**,571**	**,626**	,164	,095	-,137	-,184	,042	,104	,119	,041	-,025	,008	-,147	,066	-,373
NEO PI-R-A2	**,682**	-,103	-,347	,226	-,124	-,144	-,230	,033	-,150	-,190	,104	-,220	,341	-,141	-,096
NEO PI-R-A3	**,567**	,094	,308	-,017	,422	-,383	-,265	-,298	,023	-,014	,148	-,153	-,185	,001	,103
NEO PI-R-A4	**,649**	,141	-,362	-,273	-,068	,117	-,112	-,259	,150	-,317	-,059	,342	-,044	-,100	,000
NEO PI-R-A5	**,508**	-,376	-,433	,215	-,071	-,085	,366	-,047	,210	,025	,359	,031	-,051	,200	,036
NEO PI-R-A6	**,532**	-,091	,138	,394	**,555**	,089	,294	,112	-,167	-,186	-,193	,138	,024	,001	-,019

Hypotheses testing

Next, we computed principal component scores for each variable of the study and for each of the participants, and these scores were used as predictors in three types of regressions. First, we carried out bivariate regressions, where only the first unrotated principal component, corresponding to the BF level factor, was used as a predictor. Bivariate regressions allowed us to calculate factor models, which meant we could state whether the hypothesized factors were significant predictors of academic performance criteria. Secondly, we performed multiple regressions, where only facets (the remaining components) were introduced as predictors. These regressions permitted us to test facet models, based only on principal components representing the unique facet variance. The facet models were studied to verify whether the unique facet variance was a significant predictor of criteria of interest. Third, we ran hierarchical regressions, where in the first step the first component from a PCA (representing the corresponding BF factor) was introduced, and in the second step, the rest of the obtained components (representing facets) were entered. This type of regression allowed us to test complete "factor plus facets models", containing both factor and facet information. A comparison of the factor models with the corresponding factor plus facets models enabled us to decide whether adding the unique facet variance increased prediction significantly. Moreover, in the case of multiple and of hierarchical regressions, the squared population multiple correlations ($\hat{R}^2$) were corrected using the Browne-Cattin formula (Cattin, 1980a, 1980b). Tables 18-22 contain the [*multiple*] correlation coefficients (R) for all tested models, coefficients of determination (R^2) for factor models, and the population cross-validated multiple correlations ($\hat{R}_{cv}{}^2$) for facet models, and factors plus facets models. Next, the results of all the regressions were examined in order to test our initial hypotheses (see Chapter 2). All statistical tests were evaluated at the p = .05 Type I error rate. Below we present our main findings.

Table 18. Summary of regression coefficients of C and facets for predicting academic performance criteria.

	$R_{(D)}$	$R^2_{(D)}$	$R_{(Fs)}$	$\hat{R}_{CV}^2{}_{(Fs)}$	$R_{(D+Fs)}$	$\hat{R}_{CV}^2{}_{(D+Fs)}$	ΔR^2	ΔR^2 Average
GPA-I	,329	,108	,134	,018	,347	,121	,013	**,025**
GPA-U	,332	,110	,170	,029	,367	,135	,025	
DEV	,426	,180	,234	,055	,489	,240	,060	
GCT	,125	,016	,000	,000	,125	,016	,000	

Note: GPA-I = First year GPA standardized according to the M and SD within each college major group; GPA-U = First year GPA standardized according to general (university) M and SD; DESV = academic deviation; GCT = narrow measures of academic performance; $R_{(D)}$ = correlation coefficient for factor model; $R^2_{(D)}$ = squared sample correlation for factor model; $R_{(Fs)}$ = multiple correlation coefficient for facet model; $\hat{R}^2_{CV\,(Fs)}$ = population cross-validated multiple correlation for facet model; $R_{(D+Fs)}$ = multiple correlation coefficient for factor plus facet model; $\hat{R}^2_{CV\,(D+Fs)}$ = population cross-validated multiple correlation for factor plus facet model; ΔR^2 = the gain in R^2 when using facets in addition to the factor (calculated as $\hat{R}^2_{CV\,(D+Fs)} - R^2_{(D)}$); ΔR^2 Average = the average gain in R^2 per factor when using facets in addition to the factor.

Table 19. Summary of regression coefficients of ES and facets for predicting academic performance criteria.

	$R_{(D)}$	$R^2_{(D)}$	$R_{(Fs)}$	$\hat{R}_{CV}^{2}{}_{(Fs)}$	$R_{(D+Fs)}$	$\hat{R}_{CV}^{2}{}_{(D+Fs)}$	ΔR^2	ΔR^2 Average
GPA-I	,052	,003	,073	,005	,076	,006	,003	**,041**
GPA-U	,014	,000	,121	,015	,111	,012	,012	
DEV	,001	,000	,389	,151	,387	,150	,150	
GCT	,004	,000	,000	,000	,000	,000	,000	

Note: GPA-I = First year GPA standardized according to the M and SD within each college major group; GPA-U = First year GPA standardized according to general (university) M and SD; DESV = academic deviation; GCT = narrow measures of academic performance; $R_{(D)}$ = correlation coefficient for factor model; $R^2_{(D)}$ = squared sample correlation for factor model; $R_{(Fs)}$ = multiple correlation coefficient for facet model; $\hat{R}^2_{CV\,(Fs)}$ = population cross-validated multiple correlation for facet model; $R_{(D+Fs)}$ = multiple correlation coefficient for factor plus facet model; $\hat{R}^2_{CV\,(D+Fs)}$ = population cross-validated multiple correlation for factor plus facet model; ΔR^2 = the gain in R^2 when using facets in addition to the factor (calculated as $\hat{R}^2_{CV\,(D+Fs)} - R^2_{(D)}$); ΔR^2 Average = the average gain in R^2 per factor when using facets in addition to the factor.

Table 20. Summary of regression coefficients of E and facets for predicting academic performance criteria.

	$R_{(D)}$	$R^2_{(D)}$	$R_{(Fs)}$	$\hat{R}_{CV}^{2}{}_{(Fs)}$	$R_{(D+Fs)}$	$\hat{R}_{CV}^{2}{}_{(D+Fs)}$	ΔR^2	ΔR^2 Average
GPA-I	,061	,004	,037	,001	,040	,002	-,002	**,032**
GPA-U	,076	,006	,104	,011	,112	,012	,006	
DEV	,127	,016	,307	,094	,328	,107	,091	
GCT	,081	,007	,188	,035	,194	,038	,031	

Note: GPA-I = First year GPA standardized according to the M and SD within each college major group; GPA-U = First year GPA standardized according to general (university) M and SD; DESV = academic deviation; GCT = narrow measures of academic performance; $R_{(D)}$ = correlation coefficient for factor model; $R^2_{(D)}$ = squared sample correlation for factor model; $R_{(Fs)}$ = multiple correlation coefficient for facet model; $\hat{R}^2_{CV\,(Fs)}$ = population cross-validated multiple correlation for facet model; $R_{(D+Fs)}$ = multiple correlation coefficient for factor plus facet model; $\hat{R}^2_{CV\,(D+Fs)}$ = population cross-validated multiple correlation for factor plus facet model; ΔR^2 = the gain in R^2 when using facets in addition to the factor (calculated as $\hat{R}^2_{CV\,(D+Fs)} - R^2_{(D)}$); ΔR^2 Average = the average gain in R^2 per factor when using facets in addition to the factor.

Table 21. Summary of regression coefficients of A and facets for predicting academic performance criteria

	$R_{(D)}$	$R^2_{(D)}$	$R_{(Fs)}$	$\hat{R}^2_{CV\,(Fs)}$	$R_{(D+Fs)}$	$\hat{R}^2_{CV\,(D+Fs)}$	ΔR^2	ΔR^2 Average
GPA-I	,028	,001	,242	,058	,235	,055	,054	**,051**
GPA-U	,026	,001	,288	,083	,281	,079	,078	
DEV	,241	,058	,272	,074	,360	,130	,072	
GCT	,006	,000	,000	,000	,000	,000	,000	

Note: GPA-I = First year GPA standardized according to the M and SD within each college major group; GPA-U = First year GPA standardized according to general (university) M and SD; DESV = academic deviation; GCT = narrow measures of academic performance; $R_{(D)}$ = correlation coefficient for factor model; $R^2_{(D)}$ = squared sample correlation for factor model; $R_{(Fs)}$ = multiple correlation coefficient for facet model; $\hat{R}^2_{CV\,(Fs)}$ = population cross-validated multiple correlation for facet model; $R_{(D+Fs)}$ = multiple correlation coefficient for factor plus facet model; $\hat{R}^2_{CV\,(D+Fs)}$ = population cross-validated multiple correlation for factor plus facet model; ΔR^2 = the gain in R^2 when using facets in addition to the factor (calculated as $\hat{R}^2_{CV\,(D+Fs)} - R^2_{(D)}$); ΔR^2 Average = the average gain in R^2 per factor when using facets in addition to the factor.

Table 22. Summary of regression coefficients of O and facets for predicting academic performance criteria

	$R_{(D)}$	$R^2_{(D)}$	$R_{(Fs)}$	$\hat{R}_{CV}^2{}_{(Fs)}$	$R_{(D+Fs)}$	$\hat{R}_{CV}^2{}_{(D+Fs)}$	ΔR^2	ΔR^2 Average
GPA-I	,110	,012	,237	,056	,253	,064	,052	**,051**
GPA-U	,095	,009	,229	,052	,242	,058	,049	
DEV	,058	,003	,299	,089	,298	,089	,086	
GCT	,020	,000	,131	,017	,121	,015	,015	

Note: GPA-I = First year GPA standardized according to the M and SD within each college major group; GPA-U = First year GPA standardized according to general (university) M and SD; DESV = academic deviation; GCT = narrow measures of academic performance; $R_{(D)}$ = correlation coefficient for factor model; $R^2_{(D)}$ = squared sample correlation for factor model; $R_{(Fs)}$ = multiple correlation coefficient for facet model; $\hat{R}^2_{CV\,(Fs)}$ = population cross-validated multiple correlation for facet model; $R_{(D+Fs)}$ = multiple correlation coefficient for factor plus facet model; $\hat{R}^2_{CV\,(D+Fs)}$ = population cross-validated multiple correlation for factor plus facet model; ΔR^2 = the gain in R^2 when using facets in addition to the factor (calculated as $\hat{R}^2_{CV\,(D+Fs)} - R^2_{(D)}$); ΔR^2 Average = the average gain in R^2 per factor when using facets in addition to the factor.

Factors vs. facets as predictors of GPA. In line with Hypothesis 1, a bivariate linear regression analysis revealed that C was a significant predictor of both GPA indicators. For GPA-I, β was .33 ($p < .000$), and the factor model accounted for approximately 11 % of variance (see Table 18), $F(1) = 48.62$, $p < .000$. For GPA-U, β was .33 ($p < .000$). Conscientiousness equally explained almost 11% of variance, $F(1) = 49.42$, $p < .000$. Out of the rest of the factors, only O turned out to predict significantly one of the GPA indicators, namely GPA-I ($β = .04$, $p = .027$); explaining however, only a trivial proportion of the variance ($R^2 = .01$, $F(1) = 4.98$, $p = .027$). Remaining factors were non-significant predictors of grades. Therefore, Hypothesis 1 was strongly confirmed.

A multiple regression on 14 C facets (PC2 to PC15) revealed, that they accounted for around 4% of variance in GPA-I, $F(14) = 2.09$, $p = .012$. This percentage decreased to less than 2%, when regression result was corrected by the Browne-Cattin formula (see Table 18). Observing regression analysis summary for C facets presented in Table 23 we may conclude, that their positive relationship with GPA-I is probably due to PC4 ($β = -.13$, $p < .008$) and PC5 ($β = -.14$, $p < .005$). In the case of GPA-U, the facet model for C accounted for 5% of variance in this criterion, $F(14) = 2.51$, $p < .002$. This percentage was around 3% when the Browne-Cattin formula was applied (see Table 18). Regression coefficients for the facet model for C and GPA-U displayed in Table 24 suggest that this relationship may come from components: PC4 ($β = -.13$, $p < .007$), PC5 ($β = -.14$, $p < .004$) and PC10 ($β = .11$, $p = .024$). Therefore, Hypothesis 2 was supported.

Table 23. Regression analysis summary for C facets predicting GPA-I.

Variable	β	p
PC2	,080	,102
PC3	,080	,104
PC4	-,131	,008
PC5	-,138	,005
PC6	-,064	,190
PC7	-,036	,469
PC8	-,053	,283
PC9	,068	,167
PC10	,064	,192
PC11	-,001	,991
PC12	,039	,430
PC13	-,053	,279
PC14	,000	,993
PC15	,006	,909

Note: $\hat{R}^2$ = .037 (N = 402, p = .012), SEB = .048.

Table 24. Regression analysis summary for C facets predicting GPA-U.

Variable	β	p
PC2	,078	,109
PC3	,094	,054
PC4	-,133	,007
PC5	-,142	,004
PC6	-,093	,057
PC7	-,029	,552
PC8	-,039	,428
PC9	,060	,219
PC10	,111	,024
PC11	,013	,787
PC12	,044	,370
PC13	-,025	,607
PC14	,018	,710
PC15	,016	,748

Note: $\hat{R}^2$ = .050 (N = 402, p < .002), SEB = .049.

Likewise, when comparing R^2 found for C with $\hat{R}_{cv}^2$ found for C facets, the C factor resulted to be a much better predictor of GPA than C facets, because it explained between four and five times as much variance in grades as facets. For GPA-I this difference was 11% vs. 2%, and for GPA-U, 11% vs. 3%. These results also strongly supported Hypothesis 3.

Finally, in order to test Hypothesis 4, we compared $\hat{R}_{cv}^2$ obtained by means of hierarchical regression, where both C and C facets were used as predictors (see Table 18) and R^2, where only C was used for prediction of grades. The difference between both R^2 was of .013 in the case of GPA-I and of .025 in the case of GPA-U. In both cases, using the unique variance of the C facets would only slightly improve prediction. Therefore, Hypothesis 4 was confirmed.

Factors vs. facets as predictors of academic deviance. As revealed by the bivariate regressions for factors C and A, both factors were significant and negative predictors of academic deviance test score, $\beta = -.43$, $p < .000$ and $\beta = -.24$, $p < .000$ respectively. They also explained a significant proportion of the variance in academic deviance, in the case of C, $R^2 = .18$ (see Table 18), $F(1) = 89.36$, $p < .000$, and in the case of A, $R^2 = .06$, $F(1) = 24.87$, $p < .000$ (see Table 21). On the other hand, the results of the bivariate regression for ES showed that this factor is not significantly related to academic deviance, nor did this factor explain a significant proportion of variance. Hence, out of the three factors we expected to be significant and negative predictors of this criterion, only two, namely C and A, met our expectations. For that reason, Hypothesis 5 was only partially confirmed.

To test Hypothesis 6, the results of the academic deviance test were regressed in three separate multiple regression analyses on the facets of C, of A and of ES. The predictive model based on the C facets (PC2 to PC15) was significant, $F(14) = 3.45$, $p < .000$ and predicted 8% of variance. This percentage decreased to 5,5% when the Browne-Cattin formula for cross-validation was applied (see Table 18). From Table 25 we may infer that this significant relationship is produced by PC2 ($\beta = -.18$, $p < .000$), PC6 ($\beta = .10$, $p = .036$) and PC13 ($\beta = -.13$, $p < .006$). At the same time, the facet model

for A (PC2 to PC15) explained almost 10% of variance, $F(14) = 4.16$, $p < .000$, 7,4% when corrected by the Browne-Cattin formula (see Table 21). A detailed analysis of parameters of the facet model for A (see Table 26) indicates, that the components responsible for this relationship might be: PC2 ($\beta = .10$, $p = .041$), PC3 ($\beta = .12$, $p = .014$), PC10 ($\beta = .28$, $p < .000$) and PC14 ($\beta = .09$, $p = .049$). Finally, the facet model for ES (PC2 to PC17) predicted 18% of variance, 15% when cross-validation was applied (see Table 19) and was significant, $F(16) = 6.50$, $p < .000$. Table 27 contains the regression coefficients for ES facets. In this case, we may suppose that the significant relationship between ES facets and academic deviance is caused by the following components: PC2 ($\beta = -.14$, $p < .002$), PC4 ($\beta = -.30$, $p < .000$), PC5 ($\beta = -.22$, $p < .000$) and PC8 ($\beta = -.12$, $p < .009$). To recapitulate, the facet models for C and A and for ES all showed statistical significance, thus Hypothesis 6 was confirmed.

Table 25. Regression analysis summary for C facets predicting academic deviance.

Variable	β	p
PC2	-,183	,000
PC3	-,065	,171
PC4	-,064	,182
PC5	,047	,322
PC6	,101	,036
PC7	-,154	,001
PC8	,007	,881
PC9	-,019	,693
PC10	-,008	,874
PC11	,077	,106
PC12	,044	,361
PC13	-,131	,006
PC14	,079	,098
PC15	,022	,651

Note: $\hat{R}^2 = .078$ (N = 406, $p < .000$), SEB = .506

Table 26. Regression analysis summary for A facets predicting academic deviance.

Variable	β	p
PC2	,097	,041
PC3	,117	,014
PC4	,000	,997
PC5	,036	,452
PC6	,089	,059
PC7	,036	,447
PC8	-,083	,079
PC9	-,013	,776
PC10	,276	,000
PC11	,025	,602
PC12	,003	,951
PC13	,014	,771
PC14	,093	,049
PC15	-,058	,218

Note: $\hat{R}^2 = .098$ ($N = 406$, $p < .000$), SEB = .500

Table 27. Regression analysis summary for N/ES facets predicting academic deviance.

Variable	β	p
PC2	-,141	,002
PC3	,058	,202
PC4	-,298	,000
PC5	-,224	,000
PC6	-,042	,352
PC7	,071	,117
PC8	-,118	,009
PC9	,088	,052
PC10	-,028	,539
PC11	,035	,433
PC12	,030	,511
PC13	-,005	,912
PC14	,011	,806
PC15	,082	,069
PC16	,068	,134
PC17	,078	,083

Note: $\hat{R}^2 = .179$ ($N = 406$, $p < .000$), SEB = .477

In line with Hypothesis 7, we found support for a positive and significant relationship between E and academic deviance (β = .13, p = .01). This model explained a rather small but significant proportion of variance, $R^2 = .02$, $F(1) = 6.55$, $p = .01$.

The facet model for E explained 12% of variance in academic deviance test score, and this model was also found to be significant, $F(15) = 4.68$, $p < .000$. After the Browne-Cattin correction, this result was reduced to 9.4% of the explained variance (see Table 20). The parameters of this model may be seen in Table 28. Detailed analysis suggests, that the significant relationship between the unique variance of the E facets and academic deviance occurs due to the following components: PC3 (β = -.12, p < .006), PC4 (β = .17, p < .000), PC5 (β = .10, p = .027), PC6 (β = .14, p < .002), PC8 (β = -.14, p < .003), PC11 (β = .09, p = .044) and PC13 (β = .-17, p < .000). Therefore, Hypothesis 8 was also supported.

Table 28. Regression analysis summary for E facets predicting academic deviance.

Variable	β	p
PC2	,086	,067
PC3	-,128	,006
PC4	,174	,000
PC5	,104	,027
PC6	,145	,002
PC7	,058	,218
PC8	-,141	,003
PC9	,018	,704
PC10	-,031	,505
PC11	,094	,044
PC12	,023	,626
PC13	-,175	,000
PC14	,020	,675
PC15	,035	,454
PC16	-,033	,481

Note: $\hat{R}^2 = .120$ (N = 406, p < .000), SEB = .494

The results of comparisons between the percentage of variance explained by factors C, A, ES and E, and the percentage of variance explained by the corresponding facet models are difficult to interpret. For example, C explained 18% of variance, whereas C facets explained only 5.5%. In this case, the factor clearly was a better

predictor of academic deviance test score than the corresponding facets. However, in the case of E and A, the variance of components not related to the factor explained a bigger proportion of the criterion variance than the variance of the corresponding BF factors. While E accounted for 1,6% of academic deviance variance, the variance of components unrelated to this factor accounted for 9.4%. At the same time, while A explained approximately 6% of the variance, components not related to the A factor explained 7.4%. In turn, ES was not a significant predictor of academic deviance test score, whereas components not related to this factor accounted for a significant proportion of variance, namely 15,1%. Therefore, Hypothesis 9 was supported only for the first of the hypothesized relationships (C was a better predictor of academic deviance than the variance unrelated to this factor and corresponding to facets), while in the rest of the cases it was not confirmed (the unique variance of the A, ES and E facets accounted for a higher percentage of variance in the criterion than the corresponding personality factors). It is worth emphasizing, however, that these percentages of variance still contain an error component which could not be calculated in our study. Error variance is a part of the unique facet variance (Crocker & Algina, 1986) and is unexplainable or even random. It is possible that the variance explained by facets in the case of academic deviance criterion is mostly due to this error component. As we could not estimate this error, our results are difficult to explain. Moreover, what we related to the criterion were not personality scales *per se,* but the orthogonal components resulting from a PCA, representing the unique facet variance. These components were created for analytical purposes and cannot be directly interpreted. Therefore, for any further conclusions concerning our findings we would have to analyze the content of each of these components to identify scales that could have been a cause of these results. This is not, however, an objective of this dissertation and we must therefore leave this problem to future investigations.

Finally, to test Hypothesis 10, we carried out comparisons between $\hat{R}_{CV}^2$ obtained by means of four hierarchical regressions, where the C, A, ES and E factors, as well as their corresponding facets, were used as predictors, and R^2 resulting from bivariate regressions, where only the respective personality factors were used for academic deviance prediction (see Tables 18-21). These results demonstrated that in the case of C, the difference between both parameters was .06 in favor of the first

model. However, considering that when comparing the facet and the factor model separately facets did not predict better than factor, we may consider that this still relatively small difference in favor of facets is due to the increase in the number of predictors in regression. For A this difference was .072, for ES it was .15 and for E .091. Moreover, we previously found that the facet models for these factors predicted better than factor models. Therefore, in the case of A, ES and E the differences in favor of models including the variance not related to the corresponding BF factor were considerable. For that reason, Hypothesis 10 was supported only for C, while for A, ES and E it was not supported. Once again, we cannot explain these results with the data that we have at our disposal. Therefore, we would appeal to future investigations to further explore this matter.

Factors vs. facets as predictors of narrow measures of academic performance. The results of the bivariate regressions for BF personality factors showed that only C was a significant predictor of narrow measures of academic performance, $\beta = .12$, $p = .011$, and explained a significant, although rather small, proportion of variance in GCT scores, $R^2 = .016$, $F(1) = 6.46$, $p = .011$. Emotional stability, contrary to what we expected and similar to the rest of the factors, was found to be a non-significant predictor of GCT scores, $R^2 < .000$, $F(1) = 0.01$, *ns*. Hypothesis 11 was therefore only partially confirmed.

At the same time, the facet model for C did not explain a significant proportion of variance in narrow measures of academic performance $\hat{R}^2 < .000$, $F(14) = 0.80$, *ns*. As can be seen in Table 29 which contains the parameters of this model, none of 14 regressed components was a significant predictor of GCT scores. The facet model for ES was not statistically significant either, $\hat{R}^2 < .000$, $F(16) = 1.02$, *ns*. In Table 30, which contains coefficients for this regression, we may observe that all but one component were non-significant predictors of GCT sores. The only exception was PC16 ($\beta = -.10$, $p = .035$). According to these results, we did not find supporting evidence for Hypothesis 12.

Table 29. Regression analysis summary for C facets predicting GCT score.

Variable	β	p
PC2	-,018	,719
PC3	,022	,655
PC4	,001	,979
PC5	-,085	,088
PC6	,066	,188
PC7	,014	,786
PC8	-,049	,322
PC9	,085	,088
PC10	,014	,784
PC11	,011	,831
PC12	-,001	,986
PC13	-,027	,591
PC14	-,068	,173
PC15	,005	,921

Note: $\hat{R}^2 < .000$ (N = 406, ns.), SEB = .614

Table 30. Regression analysis summary for N/ES facets predicting GCT score.

Variable	β	p
PC2	,078	,119
PC3	-,011	,818
PC4	,080	,107
PC5	,005	,921
PC6	,021	,678
PC7	,014	,777
PC8	-,068	,169
PC9	-,009	,860
PC10	,000	,998
PC11	-,049	,329
PC12	,042	,398
PC13	,058	,241
PC14	,040	,416
PC15	,046	,360
PC16	-,105	,035
PC17	-,002	,960

Note: $\hat{R}^2 < .000$ (N = 406, ns.), SEB = .611

In line with the first part of Hypothesis 13, C was found to be a better predictor of narrow measures of academic performance than C facets. Conscientiousness accounted for a small but significant proportion of variance, whereas the facet model was not significant. At the same time, both the ES model and ES facets were not significant. Therefore, Hypothesis 13 was only partially confirmed.

Finally, in order to test Hypothesis 14, we calculated the predictive gain when facets were used in addition to factors. For C, the value of R^2 resulting from a bivariate regression was subtracted from the value of $\hat{R}_{CV}^2$ obtained by means of a hierarchical regression, where both C and C facets were used as predictors. The results showed no gain (.000) (see Table 18). For ES, we found no impact (.000) as well (see Table 19). Therefore, Hypothesis 14 was confirmed. However, the result for ES facets cannot be generalized, because both the ES factor model and the facet model for ES were statistically non-significant.

CHAPTER 4

DISCUSSION AND CONCLUSIONS

Research overview

This dissertation had two objectives. On the one hand we wanted to test whether personality factors or facets are better predictors of academic performance criteria. On the other hand, we wanted to study whether the bandwidth of the criterion to be predicted influences the predictive properties of factors and facets. We hypothesized that factors would be better predictors of academic performance outcomes regardless of their bandwidth, due to the fact that factors are latent variables underlying personality scales and not composites of facets. It is important to remember that the variance of personality measures should be considered in two parts: common and unique variance, the latter further divided into specific and error variance. Common variance is the variance shared among personality scales, and this is the part which is due to the underlying latent variables, i.e. factors. Unique variance is this part of the variance which latent variables are not able to explain and it is composed of two elements: specific variance, i.e. the true score of a facet not related to any other facet, and error variance, i.e. remaining or residual variance. Therefore, theoretically speaking, personality factors represent the common variance of personality scales, excluding the unique variance.

To achieve the goals of our investigation, we first reviewed the existing literature on bandwidth-fidelity issues within W/O psychology and we identified some limitations of previous studies. We observed that, up to the present, none of the studies have taken into account the fact that in order to estimate correctly the predictive properties of personality facets one must first separate out the variance which is due to the underlying factors. In all previous investigations, facets were treated as psychological constructs independent of the factors and no distinctions between the unique and the common facet variance were made. However, this approach is not appropriate. In fact, as mentioned before, quite the opposite occurs: facets are determined by the factor and not the other way around. A factor is not a simple sum of individual facets, but a latent variable that transfers its variance to the facets, causing their interrelationship. Therefore, a big part of the criterion variance explained by each facet, is due to one or more underlying latent variables (factors). Consequently, when analyzing the predictive properties of personality scales we must keep in mind the complex composition of their variance and treat these parts of the variance separately, something that has not been done up to now.

At the same time, we observed that comparisons of results among different studies concerning bandwidth-fidelity issues are rather difficult, as different taxonomies of personality facets are used in each case. Moreover, it is still very common for authors to refer to the same personality facet with different labels when alluding to the same personality construct or that quite the reverse happens, they use a common denomination when referring to quite distinct personality traits. Therefore, to come up with generalizable conclusions in relation to bandwidth-fidelity issues, a universal taxonomy of facets or some other reasonable alternative has to be used. We especially recommend Saucier and Ostendorf's (1999) 18 facet taxonomy which has been replicated both in English and German. We also included in our study a personality inventory based on this taxonomy and validated in Spanish, namely the B5FI (Salgado, 2006). However, currently there is no overwhelming support for any taxonomy of facets, and an alternative approach needed to be used to overcome this limitation.

The methodology of our study was designed in order to avoid the abovementioned constraints. We did not base our research on any specific taxonomy

of facets. Instead, we analyzed facets in mathematical terms, as the variance remaining after having removed factor variance from a set of personality scales. For that reason, our facets do not necessarily represent distinguishable or nameable psychological constructs.

To attain these mathematically defined facets we carried out a PCA on personality scales pertaining to three different personality inventories based on the FFM model, namely the NEO PI-R (Costa & McCrae, 1992b), the IP/5F (Salgado, 1998b), and the B5FI (Salgado, 2006). Although each of these inventories was based on a different number of personality scales, a PCA permitted us to focus on the unique variance of facets and not on a particular taxonomy of facets. The first unrotated principal component resulting from a PCA of scales pertaining to the same factor served in our study as a measure of factor. At the same time, to determine the *maximum* predictive efficiency of facets, we used the remaining unrotated principal components as the unique facet variance. This procedure permitted us to represent fully those portions of variance that do not pertain to the factor level. It also gave facets an advantage over the factor, as all their variance was taken into account, regardless of whether it was significant or not. We must mention here that exactly the same methodology was successfully used in previous investigations to separate out the specific variance from the *g* factor variance in cognitive ability measures (e.g., Ree & Earles, 1991a; Ree, Earles, & Teachout, 1994).

As a next step, after having obtained unrotated principal components for each of the personality dimensions, we estimated the principal component scores of all the study variables and for each study participant. Afterwards, we carried out three types of regression: (a) bivariate linear regressions where only factor variance was included, (b) multiple regressions where only the unique facet variance was included and (c) hierarchical regressions where, in the first step, factor variance was introduced and in the second step facet variance. Last, by means of the Browne-Cattin formula (Cattin, 1980a, 1980b) we calculated cross-validated multiple R for multiple and hierarchical regressions, so decreasing a possible bias in results which were based on multiple predictors.

At the same time, in order to analyze the impact of criterion bandwidth on the predictive properties of personality factors and facets, we used three academic

outcomes of different extension as criteria. The first criterion was GPA defined as academic grades achieved at the end of the first year of the degree. This data proceeded from academic registers and it was the broadest of all our criteria. Grade Point Average reflects a student's ability to pass successfully exams and other formal requirements of the degree. In work settings, GPA could be compared with a job performance evaluation. This criterion is also the most common one in research on academic performance. At the same time, as a criterion of the intermediate bandwidth we chose academic deviance. To measure this academic outcome we used a self-report questionnaire where students had to evaluate themselves on a 5-point scale with regard to a series of deviant academic behaviors including cheating, skipping classes, plagiarizing, etc. Lucas and Friedrich (2005) suggested that if we consider the role of a *student* as a type of *job*, academic deviance could be considered a sort of organizational deviance. From this point of view, academic deviance would be a facet of a broader construct of general academic performance. Therefore, it is undoubtedly a narrower academic performance criterion than GPA. Finally, we also used two GCT tests combined into one score as the narrowest measure of academic performance. These tests comprised two time-limited administrative tasks consisting of finding differences between two texts and of assigning people from a list to their mail box number, according to first letters of the person's last name. This type of measure focuses on specific behavior within a narrow context, and therefore is a narrower criterion than both the previous ones.

A summary of all the hypotheses tested in our study is presented in Table 31. This table illustrates that the most striking results were found for C. For all academic performance criteria the hypothesized predictive superiority of C factor over C facets was confirmed. Also hypotheses related to the lack of a significant predictive gain when using the unique variance of the C facets in addition to the factor variance for the prediction of academic performance outcomes of different bandwidths, were confirmed. In Table 18, we may observe that the percentage of variance added by C facets to the variance explained by the factor model ranged from 0% for GCT to 6% for academic deviance, and the mean was 2.5%.

In general terms, the average predictive gain across all criteria when using facets, in addition to the corresponding factors, was only 4% (see Tables 18-22). This

means that overall difference in favor of facets was trivial. To be exact, this difference was .025 for C, .041 for ES, .032 for E, .051 for A, and .051 for O. These values are of little practical importance, especially in the case of C which was the only personality dimension which predicted significantly all the academic performance criteria used in our study. On the other hand, the biggest difference in favor of facets occurred in the case of factors A and O (5.1%). However, this value may still be considered as relatively minor and irrelevant predictive gain.

Table 31. Summary of hypotheses testing.

Factor		GPA-I	GPA-U	DEV	GCT
C	Factor predicts	H1 supported	H1 supported	H5 supported	H11 supported
	Facets predict	H2 supported	H2 supported	H6 supported	H12 not supported
	Factor predicts better than facets	H3 supported	H3 supported	H9 supported	H13 supported
	Facets do not add variance	H4 supported	H4 supported	H10 supported	H14 supported
ES	Factor predicts			H5 not supported	H11 not supported
	Facets predict			H6 supported	H12 not supported
	Factor predicts better than facets			H9 not supported	H13 not supported
	Facets do not add variance			H10 not supported	H14 supported
E	Factor predicts			H7 supported	
	Facets predict			H8 supported	
	Factor predicts better than facets			H9 not supported	
	Facets do not add variance			H10 not supported	
A	Factor predicts			H5 supported	
	Facets predict			H6 supported	
	Factor predicts better than facets			H9 not supported	
	Facets do not add variance			H10 not supported	

Conclusions of the study

In summary, this dissertation made a series of contributions to W/O Psychology and to the bandwidth-fidelity debate in particular. The main conclusions are:

(1) Factors are equally good or even better predictors of academic performance criteria than facets.

(2) Adding facet variance to factor variance for the prediction of academic outcomes produces an irrelevant gain in the variance explained.

(3) Factors are the best predictors of academic performance outcomes, regardless of the bandwidth of the criterion. Our results show that the bandwidth of the academic performance criteria to be predicted does not affect the predictive capacity of personality factors. Therefore, there is no need to align the bandwidth of the predictor with the bandwidth of the criterion, as has been suggested by some authors.

(4) The results of previous studies showing the predictive superiority of facets are incorrect. Also the results of previous studies showing the predictive superiority of factors are incorrect in the choice of methods used to attain these results. The only approach which provides a definitive proof of the predictive superiority of BF factors over corresponding facets is to analyze separately the common and the unique parts of the variance of personality scales.

(5) Of all personality factors, C is the best predictor of academic outcomes, regardless of whether these outcomes are of broad, intermediate or narrow bandwidths.

With regard to the first conclusion, both for the criterion of GPA and for narrow measures of academic performance, factor variance proved to be a much better predictor than facet variance. For example in the case of grades, the unique variance of C facets explained between 1.8% and 2.9% of variance respectively for each of the GPA indicators, that is five and three less times respectively than the C factor did. For academic deviance, the C facet model also predicted less variance than the corresponding factor model, namely 5.5% versus 18%. In the case of the narrow measure of academic performance (GCT), the relationship of C facets to this criterion was found to be non-significant, while the C factor did account for a small but significant proportion of variance (1.6%). Only in the case of factors ES, E and A, for a criterion of academic deviance, did the results contradict our hypotheses and the unique facet variance show a higher relation with this criterion than factor variance. This finding could however be due to the type of instrument used to measure academic deviance, which was based on self-reporting. In the case of the remaining academic performance criteria, the instruments were based on objective measures, such as official transcripts of academic records or scores in a performance test. Therefore, it would be interesting and valuable for future investigations to further explore the relationship of personality factors and facets to this criterion and to examine whether other studies, carried out on different samples and on more objective measures of academic deviance, for example absenteeism registers, etc., come across similar results.

With respect to the second conclusion, we found that adding the unique facet variance to factor variance for the prediction of academic performance outcomes would produce on average a small, 4% gain in relative predictive efficiency across all criteria analyzed in our study. In the case of C, the best predictor of academic outcomes, these values are even smaller, 2.5% on average. Moreover, it is worth highlighting that this value corresponds to the gain due to the unique facet variance which, according to the classic measurement theory (e.g., Crocker & Algina, 1986), still contains an error component. The method used for separating the common and the unique variance in our study, a PCA, did not permit us to estimate the weight of error component and to operate on facet specific variance free of this error. However, we

may be confident that if this error had been eliminated, this already trivial result in favor of facets would have further decreased. It should also be pointed out that a PCA method puts facets at an advantage with regard to the factors, as all facet variance, even irrelevant, is considered, while normally this irrelevant variance would be discarded.

Regarding the third conclusion, our findings prove the incorrectness of the argument within the bandwidth-fidelity dilemma debate that narrow personality measures predict narrow criteria better than broad personality measures, and that they add variance to factors for the prediction of broad criteria (e.g., Ashton, 1998; Christiansen, 2008; Paunonen et al., 1999; Tett & Christiansen, 2007; Tett et al., 2003). We actually found the opposite. In the case of the broadest criterion used in our study, GPA, our results indicate that although C facets did prove to be significant predictors of this criterion, they did not predict it better than the C factor. Actually, the predictive power of all the facets together was from three to five-times weaker than in the case of the factor model. Nor did they add a significant proportion of variance above the variance explained by C factor (only 1.3% for GPA-I and 2.5% for GPA-U). At the same time, in the case of the narrowest criterion used in our study, GCT, both C and ES facets were found to be non significant predictors, while the C factor explained a small but significant proportion of variance (namely 1.6%).

Also the position within bandwidth-fidelity dilemma debate, which maintains that factors are better predictors of global criteria of performance than facets while facets are better predictors of narrow performance criteria (e.g., J. Hogan & Roberts, 1996; Moberg, 1998; Schneider et al., 1996) was not supported by our results. We found that the C factor was the best predictor both of broad GPA criterion and of the narrow measure of academic performance. It is true that in the case of academic deviance, an intermediate-bandwidth criterion, results are somewhat difficult to interpret, however, in general terms, unique facet variance was not needed to predict academic outcomes, regardless of their bandwidth. Using only factor scores, without taking into account the unique facet variance, would result in a trivial loss in predictive efficiency, namely 4%.

Concerning the fourth conclusion, our findings seem to support the position of authors who assume that broad personality measures predict broad criteria better

than narrow personality measures and, at the same time, predict equally well narrow criteria (e.g., B. P. Chapman, 2007; Mount & Barrick, 1995; Ones & Viswesvaran, 1996). As was already mentioned, we found that the average incremental value of personality facets was small for all criteria and reached a mere 4%. Therefore, we also concluded that broad personality measures not only predicted broad criteria better than facets, but they also predicted narrow criteria equally well or in some cases even better than facets. However, even if our results are in line with the position of the abovementioned authors, we must highlight that they did not use adequate analytical methods to come up with the same conclusion. The identical observation can also be made with regard to studies the results of which contradicted our findings. These studies established the predictive superiority of facets over factors; however, their methodology was also inadequate. The only correct way to definitively conclude whether facets are better predictors of any criterion than factors, or that the reverse is true, is to consider only the part of facet variance which is unrelated to factor variance. As far as we are currently aware, our study is the first one to separate the unique and the common variance in order to estimate the real predictive properties of personality factors and facets.

With regard to the fifth conclusion, our results indicate that out of all personality factors, only C resulted to be a valid and significant predictor of all outcome variables used in our study, namely: GPA, academic deviance and narrow measure of academic performance (GCT). Its predictive validity ranged from .42 for academic deviance through .33 for both GPA indicators, to .12 for GCT. Overall, C explained 18% of variance in academic deviance, 11% percent of variance in GPA, and 1.6% of variance in GCT results. As far as GPA is concerned, our results are in line with meta-analytical studies on academic performance, in finding C to be the best predictor of academic records (O'Connor & Paunonen, 2007; Poropat, 2005, 2009; Salgado, 2000; Trapmann et al., 2007). However, in the present dissertation we found in addition that C also predicts efficiently narrower than GPA academic performance criteria, namely academic deviance and narrow measures of academic performance.

Implications for future research and W/O practice

These findings may have implications both for future investigations and for W/O practice. When taking crucial decisions on the methodology to follow when personality measures are involved, researchers face a dilemma: Which personality level would be the most appropriate to predict the criteria of interest, broad or specific? At the same time, professionals must deal with a different but equally relevant concern: Is my investment in measuring facets in addition to factors profitable? This is due to the fact that measuring both factor and facets in applied settings implies additional resources (e.g., time, personnel) the financial cost of which can be easily calculated. Equally, in academic settings where personality measures are used for the admission of students, very often there exists a dilemma over whether to choose singular personality traits or general personality factors as predictors.

Due to the fact that some researchers have suggested that facets relate to many criteria in a different way than the corresponding factors, additional resources were invested in identifying job- and academic-related personality traits crucial for different organizational purposes, amongst others, selection/admission processes. However, if our findings for academic performance criteria could also be extrapolated to other than academic settings, this might cause a change in the way personality factors are seen by researchers and practitioners. Therefore, in terms of a direction for future research, it would be recommendable and valuable to extend our study also to other than academic performance criteria important to W/O psychology, for example job performance, counterproductive work behaviors, training proficiency, etc., to further bolster understanding of the role of personality facets in prediction.

If the gain attained when using facets in addition to factors for predicting criteria of different bandwidth is in fact trivial, regardless of the criterion, it could be that the measurement of facets is not worth the additional resources it requires. Knowing the predictive gain when using facets in addition to factors, both researchers and practitioners can consider the purpose of their study or of test, and decide whether or not it is worthwhile to also assess personality on a facet level in their particular case.

Future studies may also extend this dissertation by using a taxonomy of facets or other methods permitting the interpretation of facets. We opted for mathematically defined facets, which allowed us to overcome the problems related to the lack of, at the present moment, a universal taxonomy of facets, and to evoke facet collinearity. Although the fact that our facets are not directly recognizable does not constitute any constraint for our study, the ability to interpret facets would be useful. For example, in the case of academic deviance, in order to interpret our results, an additional insight into the content of principal components would be necessary. This could however be avoided if nameable and classifiable facets were available.

Limitations of the study

In the end, we would like to highlight that in this dissertation we did our best to avoid possible limitations; however, we are aware that it might be argued that some improvements could be made.

For example, one could argue that if we had applied the Schmid-Leiman method (Schmid & Leiman, 1957) of HFA instead of PCA to separate out factor and facet variance, it could perhaps have provided us with some additional advantages, such as diminishing the number of facets and improving their interpretability. Nevertheless, as we could not find any statistical program which would allow us to calculate residualized weights of the second order factors (i.e. facets) resulting from this type of factor analysis, we had to opt for an alternative method. However, we must highlight that in the case of *g* factor its estimates from PCA, PFA and HFA correlated greater than .996 (Ree & Earles, 1991b), which means that these three methods give nearly identical results. At the same time, PCA as an alternative to the third remaining option, PFA, possesses a series of additional advantages. For example, PCA is mathematically less complicated and also more accessible (Ree et al., 1994). Another important advantage of principal components is that it represents fully those portions of variance which do not belong to the common underlying factor. Last but not the least, principal components have the additional benefit of being orthogonal which avoids the problems of collinearity and enhances their usefulness in regressions

(Jolliffe, 2002). Therefore, while being interchangeable, each of these methods possesses a series of advantages and disadvantages.

Another perhaps debatable choice is the fact that for a GPA criterion we used data only for the first year academic results. To some researchers it could be arguable whether GPA from the first year is preferable to overall GPA. The reason for our decision was simply the fact that we wanted to use a comparable criterion for all study participants who were at various stages of completion of their degrees. Therefore, GPA corresponding to the first year seemed to be the only reasonable solution, because all study participants had finished at least their first year satisfactorily. This choice was also supported by the high reliability found for this criterion. The internal consistency coefficients found for first-year grades corresponding to different college majors ranged from .77 to .93, with the weighted mean of .87. This value is even slightly superior to the value used for grades in the meta-analytical studies of Roth et al. (1996) and of Credé and Kuncel (2008). We also believe that GPA corresponding to the first year is a less contaminated measure of academic performance than overall GPA, and it possesses a smaller range restriction. Overall GPA possibly measures slightly different aspects of academic performance than GPA corresponding to the first-year only, as expectations with regard to the students as well as their obligations change over the duration of their degree. However, both measures are related and show similar relations with a range of criteria, amongst others with HSGPA (Kuncel & Hezlett, 2007; Robbins et al., 2006; Robbins et al., 2004) and with results of standardized scholastic aptitudes tests, e.g., SAT, ACT, GMAT and GRE (ACT, 2007; Kobrin et al., 2008; Kuncel, Hezlett, & Ones, 2001; Robbins et al., 2004). This may suggest that these criteria are highly interchangeable.

***

In summary, our study proved that in the case of academic performance criteria, the predictive gain when using facets in addition to factors is really low and practically irrelevant (barely 4%). At the same time, we found that factors are better predictors than facets of academic outcomes regardless of their bandwidths. These results were achieved using statistical procedures which separated the variance due to the corresponding factor from facet variance, avoiding therefore the methodological

limitations of previous studies on bandwidth fidelity dilemma. Our findings may have important implications for W/O Psychology research and practice, as measurement of facets in addition to factors requires additional financial cost which could be avoided, if only factors were measured. We have indicated the average predictive gain of using facets in addition to factors, offering to researchers and to scientists the possibility to calculate the profit of using one approach in comparison to the other. Therefore, our results may help both researchers and practitioners to take more objective decisions while opting for a factor or facet level of analysis in their professional activities.

REFERENCES

ACT. (2007). The ACT. Technical Manual. Retrieved November the 10th, 2009, from http://www.act.org/aap/pdf/ACT_Technical_Manual.pdf

Albrecht, D. D., Carpenter, D. S., & Sivo, S. A. (1994). The effect of college activities and grades on job placement potential. *National Association of Student Personnel Administrators Journal, 31*(4), 290-297.

Alexander, W. P. (1935). Intelligence, concrete and abstract. *British Journal of Psychology, Monographs Supplement, 19*.

Almagor, M., Tellegen, A., & Waller, N. G. (1995). The Big Seven model: A cross-cultural replication and further exploration of the basic dimensions of natural language trait descriptors. *Journal of Personality and Social Psychology, 69*(2), 300-307.

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education, 49*(1), 62-79.

Allport, G. W., & Odbert, H. S. (1936). Trait names. A psycho-lexical study. *Psychological Monographs, 47*(211), 1-171.

Anaya, G. (1999). College impact on student learning: Comparing the use of self-reported gains, standardized test scores, and college grades. *Research in Higher Education, 40*(5), 499-526.

Anderson, G. D., & Viswesvaran, C. (1998). *An update of the validity of personality scales in personnel selection: a meta-analysis of studies published between 1992-1997.* Paper presented at the 13th Annual Conference of the Society of Industrial and Organizational Psychology, Dallas.

Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality. Special Issue: Personality language, 4*(2), 89-118.

Annen, H., & Eggimann, N. (2006). *Assessment Center Result as Predictor of Study Success.* Paper presented at the International Military Testing Association 48th Conference, Kingston.

Aoki, T. (1971). A psycho-lexical study of personality trait words: Selection, classification and desirability ratings of 455 words. *Japanese Journal of Psychology, 42*(1), 1-13.

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior, 19*(3), 289-303.

Bacon, D. R., & Bean, B. (2006). GPA in Research Studies: An Invaluable but Neglected Opportunity. *Journal of Marketing Education, 28*(1), 35-42.

Baird, L. L. (1985). Do Grades and Tests Predict Adult Accomplishment. *Research in Higher Education, 23*(1), 3-85.

Baron, J., & Norman, M. F. (1992). SATs, Achievement-Tests, and High-School Class Rank as Predictors of College Performance. *Educational and Psychological Measurement, 52*(4), 1047-1055.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1-26.

Barrick, M. R., & Mount, M. K. (2003). Impact of meta-analysis methods on understanding personality-performance relations. In K. R. Murphy (Ed.), *Validity generalization: A critical review.* (pp. 197-221). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Barrick, M. R., & Mount, M. K. (2005). Yes, Personality Matters: Moving on to More Important Matters. *Human Performance, 18*(4), 359-372.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9,* 9-30.

Bassiri, D., & Schulz, E. M. (2003). Constructing a universal scale of high school course difficulty. *Journal of Educational Measurement, 40*(2), 147-161.

Baumgarten, F. (1933). Die Charaktereigenschaften (The character traits). In A. F. Bern (Ed.), *Beitraege zur Charakter-und Persoenlichkeitsforschung (Nº 1).* Bern: A Francke.

Benet-Martínez, V., & Waller, N. G. (1997). Further evidence for the cross-cultural generality of the big seven factor model: Indigenous and imported Spanish personality constructs. *Journal of Personality, 65*(3), 567-598.

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal Deviance, Organizational Deviance, and Their Common Correlates: A Review and Meta-Analysis. *Journal of Applied Psychology, 92*(2), 410-424.

Berry, J. W. (1969). On Cross-Cultural Comparability. *International Journal of Psychology, 4*(2), 119-128.

Blankenship, K. L., & Whitley, B. E. (2000). Relation of general deviance to academic dishonesty. *Ethics & Behavior, 10*(1), 1-12.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*(2), 187-215.

Block, J. (2001). Millennial contrarianism: The five-factor approach to personality description 5 years later. *Journal of Research in Personality, 35*(1), 98-107.

Boies, K., Lee, K., Ashton, M. C., Pascal, S., & Nicol, A. A. M. (2001). The structure of the French personality lexicon. *European Journal of Personality, 15*(4), 277-295.

Bono, J. E., & Judge, T. A. (2004). Personality and Transformational and Transactional Leadership: A Meta-Analysis. *Journal of Applied Psychology, 89*(5), 901-910.

Borg, M. O., & Shapiro, S. L. (1996). Personality type and student performance in principles of economics. *Journal of Economic Education, 27*(1), 3-25.

Borg, W. R. (1963). GRE Aptitude Scores as Predictors of GPA for Graduate-Students in Education. *Educational and Psychological Measurement, 23*(2), 379-382.

Borgatta, E. F. (1964a). The structure of personality characteristics. *Behavioral Science, 9*(1), 8-17.

Borgatta, E. F. (1964b). A very short test of personality: The Behavioral Self-Rating (BSR) form. *Psychological Reports, 14*(1), 275-284.

Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences, 11*(5), 515-524.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco, CA: Jossey Bass.

Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment, 9*(1-2), 52-69.

Botwin, M. D., & Buss, D. M. (1989). Structure of act-report data: Is the five-factor model of personality recaptured? *Journal of Personality and Social Psychology, 56*(6), 988-1001.

Briggs, S. R. (1989). The optimal level of measurement for personality constructs. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 246-260). New York: Springer-Verlag.

Briggs, S. R. (1992). Assessing the five-factor model of personality description. *Journal of Personality. Special Issue: The five-factor model: Issues and applications, 60*(2), 253-293.

Brody, N. (2000). History of theories and measurements of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 16-33). New York: Cambridge University Press.

Brokken, F. B. (1978). *The language of personality.* Unpublished doctoral dissertation, University of Groningen, the Netherlands.

Brown, B. K., & Campion, M. A. (1994). Biodata Phenomenology - Recruiters Perceptions and Use of Biographical Information in Resume Screening. *Journal of Applied Psychology, 79*(6), 897-908.

Browne, M. W. (1975). Predictive validity of a linear-regression equation. *British Journal of Mathematical and Statistical Psychology, 28*(2), 79-87.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences, 29*(6), 1057-1068.

Buss, A. H. (1989). Personality as traits. *American Psychologist, 44*(11), 1378-1388.

Cable, D., & Judge, T. A. (1997). Interviewers' perceptions of person-organization fit and organizational selection decisions. *Journal of Applied Psychology, 82*(4), 546-561.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687-732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1992). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco, CA: Jossey Bass.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco, CA: Jossey Bass.

Caprara, G. V., Barbaranelli, C., & Borgogni, L. (1993). *BFQ: Big Five Questionnaire. Manuale (2nd ed.)*. Firenze: O.S. Organizzazioni Speciali.

Caprara, G. V., & Perugini, M. (1994). Personality described by adjectives: The generalizability of the Big Five to the Italian lexical context. *European Journal of Personality, 8*(5), 357-369.

Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology, 38*(4), 476-506.

Cattell, R. B. (1945). The description of personality: principles and findings in a factor analysis. *American Journal of Psychology, 58*, 69-90.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika, 12*, 197-220.

Cattell, R. B. (1948). The primary personality factors in women compared with those in men. *British Journal of Psychology, 1*, 114-130.

Cattell, R. B. (1966). Scree test for number of factors. *Multivariate Behavioral Research, 1*(2), 245-276.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF)*. Champaign, IL: IPAT.

Cattin, P. (1980a). Estimation of the predictive power of a regression-model. *Journal of Applied Psychology, 65*(4), 407-414.

Cattin, P. (1980b). Note on the estimation of the squared cross-validated multiple correlation of a regression-model. *Psychological Bulletin, 87*(1), 63-65.

Clarke, S., & Robertson, I. T. (2005). A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings. *Journal of Occupational and Organizational Psychology, 78*(3), 355-376.

Coleman, V. I., & Borman, W. C. (2000). Investigating the underlying structure of the citizenship performance domain. *Human Resource Management Review, 10*(1), 25-44.

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*(5), 678-707.

Conley, J. J. (1985). Longitudinal stability of personality traits: A multitrait-multimethod-multioccasion analysis. *Journal of Personality and Social Psychology, 49*(5), 1266-1282.

Conte, J. M., & Gintoft, J. N. (2005). Polychronicity, big five personality dimensions, and sales performance. *Human Performance, 18*(4), 427-444.

Costa, P. T., & McCrae, R. R. (1976). Age differences in personality structure: A cluster analytic approach. *Journal of Gerontology, 31*(5), 564-570.

Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1988). From Catalog to Classification - Murray Needs and the 5-Factor Model. *Journal of Personality and Social Psychology, 55*(2), 258-265.

Costa, P. T., & McCrae, R. R. (1989). *NEO PI/FFI Manual Supplement.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653-665.

Costa, P. T., & McCrae, R. R. (1992b). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Costa, P. T., & McCrae, R. R. (1995). Domains and Facets - Hierarchical Personality - Assessment Using the Revised NEO Personality-Inventory. *Journal of Personality Assessment, 64*(1), 21-50.

Costa, P. T., & McCrae, R. R. (2006). Age changes in personality and their origins: Comment on Roberts, Walton, and Viechtbauer (2006). *Psychological Bulletin, 132*(1), 26-28.

Costa, P. T., & McCrae, R. R. (2008). *Inventario de Personalidad NEO Revisado (NEO PI-R). Inventario NEO reducido de Cinco Factores (NEO-FFI), Manual 3ª edición*. Madrid: TEA.

Costa, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences, 12*(9), 887-898.

Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of personality and social psychology, 81*(2), 322-331.

Costa, P. T., Yang, J., & McCrae, R. R. (1998). Aging and personality traits: Generalizations and clinical implications. In I. H. Nordhus, G. R. VandenBos, P. Fromhold & S. Berg (Eds.), *Clinical Geropsychology* (pp. 33-48). Washington, DC: American Psychological Association.

Crant, J. M. (1995). The Proactive Personality Scale and Objective Job-Performance among Real-Estate Agents. *Journal of Applied Psychology, 80*(4), 532-537.

Credé, M., & Kuncel, N. R. (2008). Study Habits, Skills, and Attitudes The Third Pillar Supporting Collegiate Academic Performance. *Perspectives on Psychological Science, 3*(6), 425-453.

Crocker, L., & Algina, J. (Eds.). (1986). *Introduction to classical and modern test theory*. New York: CBS College Publishing.

Cronbach, L. J., & Gleser, G. C. (1957). *Psychological tests and personnel decisions*. Urbana: University of Illinois Press.

Chamorro-Premuzic, T., & Furnham, A. (2003a). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality, 37*(4), 319-338.

Chamorro-Premuzic, T., & Furnham, A. (2003b). Personality traits and academic examination performance. *European Journal of Personality, 17*(3), 237-250.

Chamorro-Premuzic, T., Furnham, A., & Ackerman, P. L. (2006). Incremental validity of the typical intellectual engagement scale as predictor of different academic performance measures. *Journal of Personality Assessment, 87*(3), 261-268.

Chapman, B. P. (2007). Bandwidth and fidelity on the NEO-Five Factor Inventory: replicability and reliability of Saucier's (1998) item cluster subcomponents. *Journal of Personality Assessment, 88*(2), 220-234.

Chapman, K. J., Davis, R., Toy, D., & Wright, L. (2004). Academic integrity in the business school environment: I'll get by with a little help from my friends. *Journal of Marketing Education, 26*(3), 236-249.

Christiansen, N. D. (2008). *Further consideration of the use of narrow trait scales.* Paper presented at the 23th Annual Conference of the Society of Industrial and Organizational Psychology, San Francisco, CA.

Church, A. T., Katigbak, M. S., & Reyes, J. A. S. (1996). Toward a taxonomy of trait adjectives in Filipino: Comparing personality lexicons across cultures. *European Journal of Personality, 10*(1), 3-24.

Church, A. T., Katigbak, M. S., & Reyes, J. A. S. (1998). Further exploration of Filipino personality structure using the lexical approach: Do the Big-Five or Big-Seven dimensions emerge? *International Congress of Psychology, Aug 1996, PQ, Canada, 12*(4), 249-269.

Church, A. T., Reyes, J. A. S., Katigbak, M. S., & Grimm, S. D. (1997). Filipino personality structure and the Big Five model: A lexical approach. *Annual Convention of American Psychological Association, Aug 1995, New York, NY, US, 65*(3), 477-528.

Dalal, R. S. (2005). A Meta-Analysis of the Relationship Between Organizational Citizenship Behavior and Counterproductive Work Behavior. *Annual Meeting of the Society for Industrial and Organizational Psychology, Apr 2004, Chicago, IL, US, 90*(6), 1241-1255.

De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology, 91*(3), 538-552.

De Fruyt, F., De Bolle, M., McCrae, R. R., Terracciano, A., Costa, P. T., & Collaborators of the Adolescent Personality Profiles of Cultures Project. (2009). Assessing the Universal Structure of Personality in Early Adolescence The NEO-PI-R and NEO-PI-3 in 24 Cultures. *Assessment, 16*(3), 301-311.

De Fruyt, F., & Mervielde, I. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality. Special Issue: Personality, learning, and education, 10*(5), 405-425.

De Fruyt, F., Mervielde, I., Hoekstra, H. A., & Rolland, J. P. (2000). Assessing adolescents' personality with the NEO PI-R. *Assessment, 7*(4), 329-345.

De Raad, B. (2000). *The Big Five Personality Factors: The psycholexical approach to personality*. Ashland, OH, US: Hogrefe & Huber Publishers.

De Raad, B., Hendriks, A. J., & Hofstee, W. K. (1992). Towards a refined structure of personality traits. *European Journal of Personality, 6*(4), 301-319.

De Raad, B., Mulder, E., Kloosterman, K., & Hofstee, W. K. (1988). Personality-descriptive verbs. *European Journal of Personality. Special Issue: Personality measurement, 2*(2), 81-96.

De Raad, B., & Perugini, M. (2002). Big Five factor Assessment: Introduction. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 1-26). Ashland, OH, US: Hogrefe & Huber Publishers.

De Raad, B., Perugini, M., Hrebícková, M., & Szarota, P. (1998). Lingua franca of personality: Taxonomies and structures based on the psycholexical approach. *Journal of Cross-Cultural Psychology. Special Issue: Personality and its measurement in cross-cultural perspective, 29*(1), 212-232.

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality. Special Issue: Personality, learning, and education, 10*(5), 303-336.

De Raad, B., & Van Heck, G. L. (1994). The Fifth of the Big Five: Editorial. *European Journal of Personality, 8*(4), 225-227.

Di Blas, L., & Forzi, M. (1998). An alternative taxonomic study of personality-descriptive adjectives in the Italian language. *European Journal of Personality, 12*(2), 75-101.

Di Blas, L., & Forzi, M. (1999). Refining a descriptive structure of personality attributes in the Italian language: The abridged big three circumplex structure. *Journal of Personality and Social Psychology, 76*(3), 451-481.

Di Fabio, A., & Busoni, L. (2007). Fluid intelligence, personality traits and scholastic success: Empirical evidence in a sample of Italian high school students. *Personality and Individual Differences, 43*(8), 2095-2104.

Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology, 41*(1), 417-440.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50*(1), 116-123.

Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research, 16*(2), 149-170.

Dilchert, S., Ones, D. S., Van Rooy, D. L., & Viswesvaran, C. (2006). Big Five factors of personality. In J. H. Greenhaus & G. A. Callanan (Eds.), *Encyclopedia of career development* (pp. 36-42). Thousand Oaks, CA: Sage.

Dobson, P., Krapljan-Barr, P., & Vielba, C. (1999). An evaluation of the validity and fairness of the graduate management admissions test (GMAT) used for MBA selection in a UK business school. *International Journal of Selection and Assessment, 7*(4), 196-202.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A Meta-Analytic Investigation of Conscientiousness in the Prediction of Job Performance: Examining the Intercorrelations and the Incremental Validity of Narrow Traits. *Journal of Applied Psychology, 91*(1), 40-57.

Elshout, J. J. (1999). De Vijf Persoonlijkheidsfactoren Test (5PFT) 1973–1999. / The Five Personality Factor Test (5PFT), 1973–1999. *Nederlands tijdschrift voor de psychologie en haar grensgebieden, 54*(5), 195-207.

Elshout, J. J., & Akkerman, A. E. (1975). *Vijf Persoonlijkheidsfactoren Test 5PFT, Handleiding / Five Personality-Factor Test 5PFT, Manual*. Nijmegen, the Netherlands: Berkhout Nijmegen B.V.

Etaugh, A. F., Etaugh, C. F., & Hurd, D. E. (1972). Reliability of college grades and grade point averages: Some implications for predictions of academic performance. *Educational and Psychological Measurement, 32*(4), 1045-1050.

Eysenck, H. J. (1947). *Dimensions of personality*. Oxford, England: Kegan Paul.

Eysenck, H. J., & Eysenck, S. B. (1977). *Psychoticism as a dimension of personality*. Oxford, England: Carne & Russak.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of de Eysenck Personality Questionnaire*. London: Hodder & Stoughton.

Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and Individual Differences, 34*(7), 1225-1243.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2*(4), 290-309.

Field, D., & Millsap, R. E. (1991). Personality in advanced old age: Continuity or change? *Journals of Gerontology, 46*(6), 299-308.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology, 44*(3), 329-344.

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among Judges of Personality - Interpersonal-Relations, Similarity, and Acquaintanceship. *Journal of Personality and Social Psychology, 69*(4), 656-672.

García, O., Aluja, A., & García, L. F. (2004). Psychometric Properties of Goldberg's 50 Personality Markers for the Big Five Model: A Study in the Spanish Language. *European Journal of Psychological Assessment, 20*(4), 310-319.

Ghiselli, E. E., & Barthol, R. P. (1953). The validity of personality inventories in the selection of employees. *Journal of Applied Psychology, 37* (1), 18-20.

Goff, M., & Ackerman, P. L. (1992). Personality Intelligence Relations - Assessment of Typical Intellectual Engagement. *Journal of Educational Psychology, 84*(4), 537-552.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of Personality and Social Psychology* (Vol. 2, pp. 141-165). Beverly Hills, CA: Sage.

Goldberg, L. R. (1982). From Ace to Zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment Vol. 1* (pp. 203-234). Hillsdale, NJ: Erlbaum.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216-1229.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26-42.

Goldberg, L. R. (1993). The structure of phenotipic personality traits. *American Psychologist, 48*(1), 26-34.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7, pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Gough, H. G. (1957). *Manual for the California Psychological Inventory*. Palo Alto, California: Consulting Psychologists Press.

Gough, H. G., & Heilbrun, A. B. J. (1965). *The Adjective Check List Manual*. Palo Alto, CA: Consulting Psychologists Press.

Graham, L. D. (1991). Predicting Academic-Success of Students in a Master of Business Administration Program. *Educational and Psychological Measurement, 51*(3), 721-727.

Gray, E. K., & Watson, D. (2002). General and specific traits of personality and their relation to sleep and academic performance. *Journal of Personality, 70*(2), 177-206.

Griffin, B., & Hesketh, B. (2004). Why Openness to Experience is not a Good Predictor of Job Performance. *International Journal of Selection and Assessment, 12*(3), 243-251.

Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*(2), 135-164.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*(3), 562-575.

Harris, D. (1940). Factors affecting college grades: a review of the literature, 1930-1937. *Psychological Bulletin, 37*(3), 125-166.

Haskell, R. E. (2001). *Transfer of learning: Cognition, instruction, and reasoning*. San Diego, CA: Academic Press, Inc.

Hendriks, A. A. J., Hofstee, W. K., & De Raad, B. (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences, 27*(2), 307-325.

Hofstee, W. K., De Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology, 63*(1), 146-163.

Hogan, J., & Hogan, R. T. (1989). How to measure employee reliability. *Journal of Applied Psychology, 74*(2), 273-279.

Hogan, J., & Holland, B. (2003). Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective. *Journal of Applied Psychology, 88*(1), 100-112.

Hogan, J., & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 849-870). San Diego, CA: Academic Press.

Hogan, J., & Roberts, B. W. (1996). Issues and non-issues in the fidelity-bandwidth trade-off. *Journal of Organizational Behavior, 17*(6), 627-637.

Hogan, R. T. (1982). A socioanalytic theory of personality. In M. Page (Ed.), *Personality: current theory and research*. Nebraska Symposium on Motivation. Lincoln, Nebraska: Nebraska University Press.

Hogan, R. T. (2005). In Defense of Personality Measurement: New Wine for Old Whiners. *Human Performance, 18*(4), 331-341.

Hogan, R. T. (Ed.). (1986). *Manual for the Hogan Personality Inventory (HPI)*. Minneapolis, MN: National Computer Systems.

Hogan, R. T., & Hogan, J. (1992). *Hogan Personality Inventory Manual*. Tulsa, OK: Hogan Assessment Systems.

Hogan, R. T., & Hogan, J. (2002). The Hogan Personality Inventory. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 329-351). Ashland, OH, US: Hogrefe & Huber Publishers.

Hogan, R. T., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions - Questions and answers. *American Psychologist, 51*(5), 469-477.

Hough, L. M. (1992). The "Big Five" personality variables-construct confusion: Description versus prediction. *Human Performance, 5(1-2)*, 139-155.

Hough, L. M. (1998). The millennium for personality psychology: New horizons or good ole daze. *Applied Psychology: An International Review. Special Issue: Applied Psychology in the USA, 47*(2), 233-261.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*(5), 581-595.

Hough, L. M., & Furnham, A. (2003). Use of personality variables in work settings. In W. C. Borman, D. R. Ilgen & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology. Vol. 12* (pp. 131-169). Hoboken, NJ, US: John Wiley & Sons Inc.

Hough, L. M., Ones, D. S., & Viswesvaran, C. (1998). *Personality correlates of managerial performance constructs.* Paper presented at the 13th Annual Conference of the Society of Industrial and Organizational Psychology, Dallas.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, Detection and Amelioration of Adverse Impact in Personnel Selection Procedures: Issues, Evidence and Lessons Learned. *International Journal of Selection and Assessment, 9*(1&2), 152-194.

Hough, L. M., & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 31-88). San Francisco: Jossey-Bass.

Howarth, E. (1976). Were Cattell's personality sphere factors correctly identified in the first instance? *British Journal of Psychology, 67*(2), 213-230.

Howarth, E., & Browne, J. A. (1971). An item-factor-analysis of the 16 PF. *Personality: An International Journal, 2*(2), 117-139.

Hrebícková, M., & Ostendorf, F. (1995). Lexikální prístup k osobnosti. V: Klasifikace prídavných jmen do kategorií osobnostní deskripce. / Lexical approach to personality: V. Classification of adjectives into categories of personality description. *Ceskoslovenská Psychologie, 39*(3), 265-276.

Hrebícková, M., Urbánek, T., Cermák, I., Szarota, P., Ficková, E., & Orlická, L. (2002). The NEO Five-Factor Inventory in Czech, Polish, and Slovak contexts. In R. R.

McCrae & J. Allik (Eds.), *The Five-Factor model of personality across cultures.* New York, NY, US: Kluwer Academic/Plenum Publishers.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA, US: Sage Publications, Inc.

Hurtz, G. M., & Donovan, J. J. (1998). *Personality and job performance: Will the real 'Big Five' please stand up?* Paper presented at the 13th Annual Conference of the Society of Industrial and Organizational Psychology, Dallas.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and Job Performance: The Big Five Revisited. *Journal of Applied Psychology, 85*(6), 869-879.

Hutz, C. S., Nunes, C. H., Silveira, A. D., Jovana, S., Anton, M., & Wieczorek, L. S. (1998). O desenvolvimento de marcadores para a avaliação da personalidade no modelo dos cinco grandes fatores. / The development of the big five markers for personality assessment in Brazil. *Psicologia: Reflexão e Crítica, 11*(2).

Isaka, H. (1990). Factor analysis of trait terms in everyday Japanese language. *Personality and Individual Differences, 11*(2), 115-124.

Jackson, C. J., Levine, S. Z., Furnham, A., & Burr, N. (2002). Predictors of cheating behavior at a university: A lesson from the psychology of work. *Journal of Applied Social Psychology, 32*(5), 1031-1046.

Jackson, D. N. (1976). *Jackson Personality Inventory manual*. Port Huron, MI: Research Psychologists Press.

Jackson, D. N. (1984). *Personality Research Form manual (3rd ed.)*. Port Huron, MI: Research Psychologists.

Jackson, D. N. (1994). *Jackson Personality Inventory - Revised Manual*. Port Huron, MI: Sigma Assessment Systems, Inc.

Jang, K. L., Livesley, W. J., & Vernon, P. A. (1996). Heritability of the big five personality dimensions and their facets: A twin study. *Journal of Personality, 64*(3), 577-591.

Jenkins, M., & Griffith, R. (2004). Using Personality Constructs to Predict Performance: Narrow or Broad Bandwidth. *Journal of Business and Psychology, 19*(2), 255-269.

Ji, C. H. C. (1998). Predictive validity of the graduate record examination in education. *Psychological Reports, 82*(3), 899-904.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York, NY, US: Guilford Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory - Versions 4a and 54.* Berkeley: University of California at Berkeley, Institute of Personality and Social Research.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press.

John, O. P., & Robins, R. W. (1993). Gordon Allport father and critic of the Five-Factor Model. In K. H. Craik, R. Hogan & R. N. Wolfe (Eds.), *Fifty years of personality psychology*. New York: Plenum Press.

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461-494). New York, NY: Cambridge University Press.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research (2nd ed.).* (pp. 102-138). New York, NY, US: Guilford Press.

Jolliffe, I. T. (2002). *Principal component analysis. Second edition.* New York: Springer.

Jones, E. (2008). Predicting performance in first-semester college basic writers: Revisiting the role of self-beliefs. *Contemporary Educational Psychology, 33*(2), 209-238.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*(4), 765-780.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*(3), 530-541.

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology, 87*(4), 797-807.

Kahn, J. H., Nauta, M. M., Gailbreath, R. D., Tipps, J., & Chartrand, J. M. (2002). The utility of career and personality assessment in predicting academic progress. *Journal of Career Assessment, 10*(1), 3-23.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141-151.

Kanfer, R., Ackerman, P. L., Murtha, T., & Goff, M. (1995). Personality and intelligence in industrial and organizational psychology. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 577-602). New York: Plenum.

Katigbak, M. S., Church, A. T., & Akamine, T. X. (1996). Cross-Cultural Generalizability of Personality Dimensions: Relating Indigenous and Imported Dimensions in Two Cultures. *Journal of Personality and Social Psychology, 70*(1), 99-114.

Klehe, U. C., & Anderson, N. (2007). The moderating influence of personality and culture on social loafing in typical versus maximum performance situations. *International Journal of Selection and Assessment, 15*(2), 250-262.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average*. New York: The College Board.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63-82.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2007). A meta-analysis of the predictive validity of the Graduate Management Admission Test (GMAT) and Undergraduate Grade Point Average (UGPA) for graduate student academic performance. *Academy of Management Learning & Education, 6*(1), 51-68.

Kuncel, N. R., & Hezlett, S. A. (2007). Assessment - Standardized tests predict graduate students' success. *Science, 315*(5815), 1080-1081.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for

graduate student selection and performance. *Psychological Bulletin, 127*(1), 162-181.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic Performance, Career Potential, Creativity, and Job Performance: Can One Construct Predict Them All? *Journal of Personality and Social Psychology, 86*(1), 148-161.

Lautenschlager, G. J. (1990). Sources of Imprecision in Formula Cross-Validated Multiple Correlations. *Journal of Applied Psychology, 75*(4), 460-462.

Lawson, R. A. (2004). Is classroom cheating related to business students' propensity to cheat in the "real world"? *Journal of Business Ethics, 49*(2), 189-199.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research, 39*(2), 329-358.

Lee, K., & Ashton, M. C. (2006). Further assessment of the HEXACO personality inventory: Two new facet scales and an observer report form. *Psychological Assessment, 18*(2), 182-191.

LePine, J. A., Erez, A., & Johnson, D. E. (2002). The nature and dimensionality of organizational citizenship behavior: A critical review and meta-analysis. *Journal of Applied Psychology, 87*(1), 52-65.

Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology, 90*(3), 442-452.

Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: a five-factor model perspective. *Medical Education, 36*(11), 1050-1056.

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality Scale Validities Increase Throughout Medical School. *Journal of Applied Psychology, 94*(6), 1514-1535.

Lindley, L. D., & Borgen, F. H. (2002). Generalized self-efficacy, Holland theme self-efficacy, and academic performance. *Journal of Career Assessment, 10*(3), 301-314.

Locke, E. A., & Hulin, C. L. (1962). A review and evaluation of the validity studies of Activity Vector Analysis. *Personnel Psychology, 15*(2), 25-42.

Lounsbury, J. W., Gibson, L. W., & Hamrick, F. L. (2004). The development and validation of a personological measure of work drive. *Journal of Business and Psychology, 18*(4), 427-451.

Lounsbury, J. W., Sundstrom, E., Loveland, J. L., & Gibson, L. W. (2003). Intelligence, "Big Five" personality traits, and work drive as predictors of course grade. *Personality and Individual Differences, 35*(6), 1231-1239.

Lucas, G. M., & Friedrich, J. (2005). Individual Differences in Workplace Deviance and Integrity as Predictors of Academic Dishonesty. *Ethics & Behavior, 15*(1), 15-35.

Marcus, B., Lee, K., & Ashton, M. C. (2007). Personality dimensions explaining relationships between integrity tests and counterproductive behavior: Big five, or one in addition? *Personnel Psychology, 60*(1), 1-34.

Margrain, S. A. (1978). Student characteristics and academic performance in higher education: A review. *Research in Higher Education, 8*(2), 111-123.

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*(1), 139-157.

Matthews, G. (1989). The factor structure of the 16PF: Twelve primary and three secondary factors. *Personality and Individual Differences, 10*(9), 931-940.

Maulden, W. B. (2008). A Delineation of Conscientiousness: Toward an Adequate Taxonomy (Doctoral dissertation, The University of Memphis, 2008). *Dissertation Abstracts International: Section B: The Sciences and Engineering, Vol 69*(4-B), 2663.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality. Special Issue: The five-factor model: Issues and applications, 60*(2), 329-361.

McCloy, C. H. (1936). A factor analysis of personality traits to underlie character education. *Journal of Educational Psychology, 27*(5), 375-375.

McCrae, R. R., & Costa, P. T. (1985a). Openness to experience. In R. T. Hogan & W. H. Jones (Eds.), *Perspectives in personality* (Vol. 1, pp. 145-172). Greenwich, CT: JAI Press.

McCrae, R. R., & Costa, P. T. (1985b). Updating Norman's "adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49*(3), 710-721.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81-90.

McCrae, R. R., Costa, P. T., Terracciano, A., Parker, W. D., Mills, C. J., De Fruyt, F., et al. (2002). Personality trait development from age 12 to age 18: Longitudinal, cross-sectional, and cross-cultural analyses. *Journal of Personality and Social Psychology, 83*(6), 1456-1468.

McCrae, R. R., Terracciano, A., & 78 members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*(3), 547-561.

McCrae, R. R., Yik, M. S. M., Trapnell, P. D., Bond, M. H., & Paulhus, D. L. (1998). Interpreting personality profiles across cultures: Bilingual, acculturation, and peer rating studies of Chinese undergraduates. *Journal of Personality and Social Psychology, 74*(4), 1041-1055.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The Validity of Employment Interviews - a Comprehensive Review and Metaanalysis. *Journal of Applied Psychology, 79*(4), 599-616.

McDonald, R. P. (1981). The dimensionality of tests and ítems. *British Journal of Mathematical and Statistical Psychology, 34*(1), 100-107.

Mershon, B., & Gorsuch, R. L. (1988). Number of factors in the personality sphere: Does increase in factors increase predictability of real-life criteria? *Journal of Personality and Social Psychology, 55,* 675-680.

Mervielde, I., & De Fruyt, F. (2000). The big five personality factors as a model for the structure of children's peer nominations. *European Journal of Personality, 14*(2), 91-106.

Messick, S. (1979). Potential uses of noncognitive measurement in education. *Journal of Educational Psychology, 71*(3), 281-292.

Mischel, W. (1968). *Personality and assessment*. Hoboken, NJ, US: John Wiley & Sons Inc.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80*(4), 252-283.

Moberg, P. J. (1998). Predicting conflict strategy with personality traits: Incremental validity and the five factor model. *International Journal of Conflict Management, 9*(3), 258-285.

Mol, S. T., Born, M. P. H., Willemsen, M. E., & Van Der Molen, H. T. (2005). Predicting Expatriate Job Performance for Selection Purposes: A Quantitative Review. *Journal of Cross-Cultural Psychology, 36*(5), 590-620.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007a). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology, 60*(4), 1029-1049.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007b). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*(3), 683-729.

Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimesions: implications for research and practice in human resources management. In K.M. Rowland & G. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 13, pp. 153-200). Greenwich, CT: JAI Press Inc.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-Factor Model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*(2-3), 145-165.

Munson, H. L., & Rubenstein, B. J. (1992). School IS work: Work task learning in the classroom. *Journal of Career Development, 18*(4), 289-297.

Musgrave-Marquart, D., Bromley, S. P., & Dalley, M. B. (1997). Personality, academic attribution, and substance use as predictors of academic achievement in college students. *Journal of Social Behavior and Personality, 12*(2), 501-511.

Newman, R. I. (1968). GRE Scores as Predictors of GPA for Psychology Graduate Students. *Educational and Psychological Measurement, 28*(2), 433-436.

Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success. A meta-analysis. *Personnel Psychology, 58*(2), 367-408.

Nilsson, J. E. (1995). The GRE and the GMAT - a comparison of their correlations to GGPA. *Educational and Psychological Measurement, 55*(4), 637-640.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116-130.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology, 66*(6), 574-583.

Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population.* University of Michigan, Department of Psychology.

Nunally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.

O'Connor, M. C., & Paunonen, S. V. (2007). Big five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*(5), 971-990.

Ones, D. S., & Anderson, N. (2002). Gender and ethnic group differences on personality scales in selection: Some British data. *Journal of Occupational and Organizational Psychology, 75*, 255-276.

Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology, 60*(4), 995-1027.

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609-626.

Ones, D. S., & Viswesvaran, C. (2003). Job-specific applicant pools and national norms for personality scales: Implications for range-restriction corrections in validation research. *Journal of Applied Psychology, 88*(3), 570-577.

Ostendorf, F. (1990). *Sprache und Persoenlichkeitsstruktur: Zur Validitaet des Fuenf-Faktoren-Modells der Persoenlichkeit [Language and Personality Structure: On*

the Validity of the Five Factor Model of Personality]. Regensburg, Germany: S. Roderer Verlag.

Oswald, F. L., Schmit, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*(2), 187-207.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students. Findings and insights from twenty years of research.* San Francisco, CA: Jossey-Bass.

Paunonen, S. V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology, 74*(2), 538-556.

Paunonen, S. V. (2003). Big five factors of personality and replicated predictions of nehavior. *Journal of Personality and Social Psychology, 84*(2), 411-424.

Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality, 35*(1), 78-90.

Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? Plenty! *Journal of Personality, 68*(5), 821-835.

Paunonen, S. V., Jackson, D. N., Trzebinski, J., & Forsterling, F. (1992). Personality Structure across Cultures - a Multimethod Evaluation. *Journal of Personality and Social Psychology, 62*(3), 447-456.

Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior, 20*(3), 389-405.

Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57*(3), 552-567.

Peeters, H., & Lievens, F. (2005). Situational judgment tests and their predictiveness of college students' success: The influence of faking. *Educational and Psychological Measurement, 65*(1), 70-89.

Peeters, M. A. G., Van Tuijl, H. F. J. M., Rutte, C. G., & Reymen, I. M. M. J. (2006). Personality and Team Performance: A Meta-Analysis. *European Journal of Personality, 20*(5), 377-396.

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry, 5*(2), 103-113.

Poropat, A. E. (2002). When good guys come first: An examination of the relationship between citizenship and task performance in an academic setting. *Envisioning Practice - Implementing Change, Vol 3*, 11-18.

Poropat, A. E. (2005). *An examination of the relationship between personality and citizenship performance in academic and workplace settings.* Unpublished doctoral dissertation, Griffith University, Brisbane, Queensland, Australia.

Poropat, A. E. (2009). A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance. *Psychological Bulletin, 135*(2), 322-338.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612-624.

Quevedo-Aguado, M. P., Iraegui, A., Anivarro, E. M., & Ross, P. (1996). Linguistic descriptors of personality in the Spanish language: A first taxonomic study. *European Journal of Personality, 10*(1), 25-34.

Ree, M. J., & Earles, J. A. (1991a). Predicting Training Success - Not Much More Than G. *Personnel Psychology, 44*(2), 321-332.

Ree, M. J., & Earles, J. A. (1991b). The stability of g across different methods of estimation. *Intelligence, 15*(3), 271-278.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting Job-Performance - Not Much More Than G. *Journal of Applied Psychology, 79*(4), 518-524.

Reilly, R. R., & Warech, M. A. (1993). The validity and fairness of alternatives to cognitive tests. In L. C. Wing & B. R. Gifford (Eds.), *Policy issues in employment testing* (pp. 131-224). Boston: Kluwer.

Richter, A. (2003). Predictors of academic performance and pro-social behavior of business administration students. *Journal of Academy of Business and Economics, 2*(1), 68-86.

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality, 65*(3), 449-475.

Robbins, S. B., Allen, J., Casillas, A., Peterson, C. H., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management

measures from traditional predictors of college outcomes. *Journal of Educational Psychology, 98*(3), 598-616.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill, factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288.

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: an empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*(1), 103-139.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006a). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1-25.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006b). Personality traits change in adulthood: Reply to Costa and McCrae (2006). *Psychological Bulletin, 132*(1), 29-32.

Rode, J. C., Arthaud-Day, M. L., Mooney, C. H., Near, J. P., Baldwin, T. T., Bommer, W. H., et al. (2005). Life satisfaction and student performance. *Academy of Management Learning and Education, 4*(4), 421-433.

Rolland, J. P. (2002). The cross-cultural generalizability of the Five-Factor Model of personality. In R. R. McCrae & J. Allik (Eds.), *Five-Factor Model of Personality Across Cultures* (pp. 7-28). New York: Kluwer Academic/Plenum Publishers.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59-82.

Roth, P. L., BeVier, C. A., Switzer, F. S., & Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology, 81*(5), 548-556.

Roth, P. L., & Bobko, P. (2000). College grade point average as a personnel selection device: Ethnic group differences and potential adverse impact. *Journal of Applied Psychology, 85*(3), 399-406.

Roth, P. L., & Clarke, R. L. (1998). Meta-analyzing the relation between grades and salary. *Journal of Vocational Behavior, 53*(3), 386-400.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review, 16*(2), 155-180.

Rothstein, M. G., & Jelley, R. B. (2003). The challenge of aggregating studies of personality. In K. R. Murphy (Ed.), *Validity generalization: A critical review.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Rothstein, M. G., Paunonen, S. V., Rush, J. C., & King, G. A. (1994). Personality and Cognitive-Ability Predictors of Performance in Graduate Business School. *Journal of Educational Psychology, 86*(4), 516-530.

Rynes, S. L., Orlitzky, M. O., & Bretz, R. D. (1997). Experienced hiring versus college recruiting: Practices and emerging trends. *Personnel Psychology, 50*(2), 309-339.

Salgado, J. F. (1996). Análisis exploratorio y confirmatorio del Inventario de Personalidad de Cinco Factores (IP/5F) / Exploratory and confirmatory analysis of Five Factor Personality Inventory (IP/5F). *Psicológica, 17*, 353-366.

Salgado, J. F. (1997). The Five Factor Model of Personality and Job Performance in the European Community. *Journal of Applied Psychology, 82*(1), 30-43.

Salgado, J. F. (1998a). Big Five personality dimensions and job performance in army and civil occupations: A European perspective. *Human Performance, 11*(2-3), 271-288.

Salgado, J. F. (1998b). *Manual Técnico. Inventario de Personalidad de Cinco Factores IP5/F*. Santiago de Compostela: Tórculo Edicións, S.L.

Salgado, J. F. (2000). *The Big Five personality dimentions as predictors of alternative criteria.* Paper presented at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Salgado, J. F. (2002). The Big Five Personality Dimensions and Counterproductive Behaviors. *International Journal of Selection and Assessment, 10*(1&2), 117-125.

Salgado, J. F. (2003a). Escala de Conductas Académicas. / Academic Behaviors Scale. Unpublished manuscript. Department of General and Social Psychology and Methodology, University of Santiago de Compostela, Santiago de Compostela, Spain.

Salgado, J. F. (2003b). La hipótesis de la especifidad situacional, la hipótesis de la generalización de la validez y la predicción del desempeño ocupacional: influencia de la personalidad. Department of General and Social Psychology and Methodology, University of de Santiago de Compostela, Santiago de Compostela, Spain.

Salgado, J. F. (2003c). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology, 76*(3), 323-346.

Salgado, J. F. (2006). Development and validtion of the Big Five-Factor Inventory. Unpublished manuscript. Department of General and Social Psychology and Methodology. University of Santiago de Compostela, Santiago de Compostela, Spain.

Salgado, J. F., Moscoso, S., & Lado, M. (2003). Evidence of cross-cultural invariance of the big five personality dimensions in work settings. *European Journal of Personality, 17*, S67-S76.

Salgado, J. F., Ones, D. S., & Viswesvaran, C. (2001). Predictors used for personnel selection: An overview of constructs, methods and techniques. In N. Anderson, D. S. Ones, H. K. Sinangil & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology, Volume 1: Personnel psychology.* (pp. 165-199). Thousand Oaks, CA: Sage Publications Ltd.

Salgado, J. F., Rumbo, A., Santamaría, G., & Rodríguez-Losada, M. (1995). El 16 PF, el modelo de los cinco factores y el rendimiento en el trabajo./ 16 PF, Big five factor model and job performance. *Revista de Psicología Social y Aplicada, 5*(3), 81-94.

Samson, G. E., Graue, M. E., Weinstein, T., & Walberg, H. J. (1984). Academic and Occupational Performance - a Quantitative Synthesis. *American Educational Research Journal, 21*(2), 311-321.

Sarmiento, R., Beale, J., & Knowles, G. (2007). Determinants of performance amongst shop-floor employees. *Management Research News, 30*(12), 912-927.

Saucier, G. (1994). Trapnell versus the lexical factor: More ado about nothing? *European Journal of Personality. Special Issue: The fifth of the Big Five, 8*(4), 291-298.

Saucier, G., Georgiades, S., Tsaousis, I., & Goldberg, L. R. (2005). The factor structure of Greek personality adjectives. *Journal of Personality and Social Psychology, 88*(5), 856-875.

Saucier, G., & Goldberg, L. R. (2002). Assessing the Big Five: Applications of 10 psychometric criteria to the development of marker scales. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 29-58). Ashland, OH, US: Hogrefe & Huber Publishers.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: a cross-language replication. *Journal of Personality and Social Psychology, 76*(4), 613-627.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53-61.

Schmit, M. J., Kihm, J. A., & Robie, C. (2000). Development of a global measure of personality. *Personnel Psychology, 53*(1), 153-193.

Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37*(3), 407-422.

Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., et al. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology, 94*(6), 1479-1497.

Schmitt, N., Oswald, F. L., Kim, B. H., Imus, A., Merritt, S., Friede, A., et al. (2007). The use of background and ability profiles to predict college student outcomes. *Journal of Applied Psychology, 92*(1), 165-179.

Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broadsided by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behavior, 17*(6), 639-655.

Schuler, H., Funke, U., & Baronboldt, J. (1990). Predictive validity of school grades - a meta-analysis. *Applied Psychology - An International Review. Psychologie Appliquee - Revue Internationale, 39*(1), 89-103.

SHL. (1997). *Customer contact manual and user's guide*. Thames Ditton: SHL Ltd.

Shmelyov, A. G., & Pokhil'ko, V. I. (1993). A taxonomy-oriented study of Russian personality-trait names. *European Journal of Personality, 7*(1), 1-17.

Singleton, R. A., Jr. (2007). Collegiate alcohol consumption and academic performance. *Journal of Studies on Alcohol and Drugs, 68*(4), 548-555.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Smith, G. M. (1967). Usefulness of Peer Ratings of Personality in Educational Research. *Educational and Psychological Measurement, 27*(4, PT. 2), 967-984.

Smith, G. M. (1969). Relations between personality and smoking behavior in preadult subjects. *Journal of Consulting and Clinical Psychology, 33*(6), 710-715.

Smith, R. M., & Schumacher, P. A. (2005). Predicting Success for Actuarial Students in Undergraduate Mathematics Courses. *College Student Journal, 39*(1), 165-177.

Somer, O., & Goldberg, L. R. (1999). The structure of Turkish trait-descriptive adjectives. *Journal of Personality and Social Psychology, 76*(3), 431-450.

Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality, 43*(1), 84-90.

Steel, P. (2007). The Nature of Procrastination: A Meta-Analytic and Theoretical Review of Quintessential Self-Regulatory Failure. *Psychological Bulletin, 133*(1), 65-94.

Stein, M. I. (1963). *Personality measures in admissions*. New York, NY, US: College Entrance Examination Board.

Sternberg, R. J., & Wagner, R. K. (1993). The g-ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science, 2*(1), 1-5.

Stewart, G. L. (1999). Trait Bandwidth and Stages of Job Performance: Assessing Differential Effects for Conscientiousness and Its Subtraits. *Journal of Applied Psychology, 84*(6), 959-968.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*(5), 401-426.

Szarota, P. (1995). Polska Lista Przymiotnikowa (PLP): Narzędzie do diagnozy Pięciu Wielkich czynników osobowości. / Polish Adjective List: An instrument to assess the five-factor model of personality. *Studia Psychologiczne, 33*(1-2), 227-256.

Szarota, P. (1996). Taxonomy of the Polish personality-descriptive adjectives of the highest frequency of use. *Polish Psychological Bulletin, 27*(4), 343-351.

Szarota, P., Ashton, M. C., & Lee, K. (2007). Taxonomy and structure of the Polish personality lexicon. *European Journal of Personality, 21*(6), 823-852.

Szirmák, Z., & De Raad, B. (1994). Taxonomy and structure of Hungarian personality traits. *European Journal of Personality, 8*(2), 95-117.

Terracciano, A., Costa, P. T., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32*(8), 999-1009.

Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: a response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, And Schmitt (2007). *Personnel Psychology, 60*(4), 967-993.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44*(4), 703-742.

Tett, R. P., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1994). Meta-analysis of personality-job performance relations: a reply to Ones, Mount, Barrick, and Hunter (1994). *Personnel Psychology, 47*(1), 157-172.

Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality-job performance relationship. *Journal of Organizational Behavior, 24*(3), 335-356.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41*(1), 1-32.

Trapmann, S., Hell, B., Hirn, J. O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift Fur Psychologie, 215*(2), 132-151.

Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology, 59*(4), 781-790.

Tross, S. A., Harper, J. P., Osher, L. W., & Kneidinger, L. M. (2000). Not just the usual cast of characteristics: Using personality to predict college performance and retention. *Journal of College Student Development, 41*(3), 323-334.

Tupes, E. C., & Christal, R. C. (1958). *Stability of personality trait rating factors obtained under diverse conditions.* (Vol. USAF WADC Technical Note No. 58-61). Lackland Air Force Base, TX: U.S. Air Force.

Tupes, E. C., & Christal, R. C. (1961). *Recurrent personality factors based on trait ratings.* (Vol. 61-97 USAF ASD Technical Report). Lackland Air Force Base, TX: U.S. Air Force.

Vasilopoulos, N. L., Cucina, J. M., & Hunter, A. E. (2007). Personality and training proficiency: Issues of bandwidth-fidelity and curvilinearity. *Journal of Occupational and Organizational Psychology, 80*(1), 109-131.

Vinchur, A. J., Schippmann, J. S., Switzer III, F. S., & Roth, P. L. (1998). A Meta-Analytic Review of Predictors of Job Performance for Salespeople. *Journal of Applied Psychology, 83*(4), 586-597.

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology, 90*(1), 108-131.

Warr, P., Bartram, D., & Martin, T. (2005). Personality and sales performance: Situational variation and interactions between traits. *International Journal of Selection and Assessment, 13*(1), 87-91.

Webb, E. (1915). *Character and Intelligence: An attempt at an exact study of caracter.* Cambrige: Cambrige, University Press.

Wechsler, D. (1943). Non-intellective factors in general intelligence. *The Journal of Abnormal and Social Psychology, 38*(1), 101-103.

Wherry, R. J. (1931). A new formula for predicting the shrinkage of the coefficient of multiple correlation. *Annals of Mathematical Statistics, 2*, 440-457.

Wiggins, J. S. (1968). Personality Structure. *Annual Review of Psychology, 19*(1), 293-350.

Wiggins, J. S. (1982). Circumplex models of interpersonal behavior in clinical psychology. In P. S. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 183-221). New York: Wiley.

Wolfe, R. N., & Johnson, S. D. (1995). Personality as a Predictor of College Performance. *Educational and Psychological Measurement, 55*(2), 177-185.

Yang, K. S., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. *Journal of Personality and Social Psychology, 58*(6), 1087-1095.

Zajacova, A., Lynch, S. M., & Espenshade, T. J. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education, 46*(6), 677-706.

Zhao, H., & Seibert, S. E. (2006). The Big Five Personality Dimensions and Entrepreneurial Status: A Meta-Analytical Review. *Journal of Applied Psychology, 91*(2), 259-271.

ÍNDICE DE TABLAS

TABLE INDEX

ÍNDICE DE FIGURAS

FIGURE INDEX